

Brittleness of Bayesian inference and new Selberg formulas

Houman Owhadi & Clint Scovel
California Institute of Technology

October 27, 2014

The incorporation of priors [30] in the Optimal Uncertainty Quantification (OUQ) framework [31] reveals brittleness in Bayesian inference; a model may share an arbitrarily large number of finite-dimensional marginals with, or be arbitrarily close (in Prokhorov or total variation metrics) to, the data-generating distribution and still make the largest possible prediction error after conditioning on an arbitrarily large number of samples. The initial purpose of this paper is to unwrap this brittleness mechanism by providing (i) a quantitative version of the Brittleness Theorem of [30] and (ii) a detailed and comprehensive analysis of its application to the revealing example of estimating the mean of a random variable on the unit interval $[0, 1]$ using priors that exactly capture the distribution of an arbitrarily large number of Hausdorff moments.

However, in doing so, we discovered that the free parameter associated with Markov and Kreĭn's canonical representations of truncated Hausdorff moments generates reproducing kernel identities corresponding to reproducing kernel Hilbert spaces of polynomials. Furthermore, these reproducing identities lead to biorthogonal systems of Selberg integral formulas.

This process of discovery appears to be generic: whereas Karlin and Shapley used Selberg's integral formula to first compute the volume of the Hausdorff moment space (the polytope defined by the first n moments of a probability measure on the interval $[0, 1]$), we observe that the computation of that volume along with higher order moments of the uniform measure on the moment space, using different finite-dimensional representations of subsets of the infinite-dimensional set of probability measures on $[0, 1]$ representing the first n moments, leads to families of equalities corresponding to classical and new Selberg identities.

2010 Mathematics Subject Classification: 62A01, 62F12, 62F15, 62G20, 62G35. 46E22, 11M36

Keywords: Bayesian inference, misspecification, robustness, uncertainty quantification, optimal uncertainty quantification, reproducing kernel Hilbert spaces (RKHS), Selberg integral formulas.

Houman Owhadi: owhadi@caltech.edu

Clint Scovel: corresponding author, clintscovel@gmail.com

Contents

1	Introduction	3
2	OUQ with Priors	5
3	Quantification of Bayesian Brittleness	9
3.1	Application to a revealing example	10
4	Volume Inequalities on the Hausdorff Moment Space	12
5	Integral Geometry of the Markov-Kreĭn Representations	14
5.1	Principal Representations	15
5.2	Canonical Representations	18
5.3	Change of variables integral representations	23
6	Mean Hausdorff Moments using the Markov-Kreĭn Representations	24
7	The Canonical Representations and Reproducing Kernel Hilbert Spaces of Polynomials	30
8	New Selberg Integral Formulas	34
9	Proofs	36
9.1	Proof of Theorem 3.2	36
9.2	Proof of Theorem 3.6	38
9.3	Proof of Lemma 4.1	40
9.4	Proof of Lemma 4.3	40
9.5	Proof of Lemma 4.4	44
9.6	Proof of Proposition 5.5	45
9.7	Proof of Proposition 5.6	47
9.8	Proof of Proposition 6.1	50
9.9	Proof of Theorem 7.1	54
9.10	Proof of Theorem 8.1	56
9.11	Proof of Theorem 8.2	58
9.12	Proof of Theorem 8.4	59
9.13	Proof of Proposition 9.2	63
10	Appendix	65
10.1	Proof of even case of Lemma 4.3	65
10.2	Even case of Proof of Theorem 7.1	68
10.3	Assorted Technical Results	69
	Acknowledgements	70
	References	70

1 Introduction

Optimal Uncertainty Quantification (OUQ) [31] provides a framework for the computation of optimal bounds on quantities of interest -given a set of available information and specified assumptions. Although this framework is neither frequentist nor Bayesian, in that it is simply expressed in terms of optimization over measures and functions, a natural question arises; what happens when we introduce priors into OUQ? In Owhadi et al. [30], this program was initiated through the introduction of a further set of assumptions, namely, the assumptions regarding the prior on the specified assumption set. A corresponding reduction theory for optimization problems over measures on spaces of measures is established, facilitating the computation of optimal bounds on prior and posterior values and the analysis of the consequences of conditioning on observed data. However, the completion of this program reveals Brittleness theorems [30, Thm. 4.13, Thm. 6.4, Thm. 6.9] for Bayesian Inference -mild assumptions are sufficient to demonstrate that, given a set of priors, conditioning on observations can produce arbitrary results, regardless of the sample size.

Although it is known from the results of Diaconis and Freedman that the Bayesian method may fail to converge or may converge towards the wrong solution (i.e., be inconsistent) if the underlying probability mechanism allows an infinite number of possible outcomes [14] and that in these non-finite-probability-space situations, this lack of convergence (commonly referred to as *Bayesian inconsistency*) is the rule rather than the exception [15], it is also known, from the Bernstein-Von Mises Theorem [7, 42] (see also LeCam [28]), that consistency (convergence upon observation of sample data) does indeed hold, under some regularity conditions, if the data-generating distribution of the sample data belongs to the finite dimensional family of distributions parameterized by the model. Furthermore, although it is also known that this convergence may fail under model misspecification [43, 21, 32, 1, 2, 26, 29, 22] (i.e. when the data-generating distribution does not belong to the family of distributions parameterized by the model), it is natural to wonder whether a “close enough” model has good convergence properties: see e.g. [17, 37, 18] and in particular G. E. P. Box’s question [?, p. 74] “Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful?”

The Brittleness theorems [30, Thm. 4.13, Thm. 6.4, Thm. 6.9] suggest that there may be no such thing as a “close enough” model if Box’s question is answered in the classical framework of Bayesian sensitivity analysis (where given the data and a class of priors one computes optimal bounds on posterior values); indeed, if “closeness” is defined (i) as sharing an arbitrarily large finite number of finite-dimensional marginals or (ii) using the Prokhorov or total variation metrics, then the posterior values of such “close” models may be as distant as possible after conditioning on an arbitrarily large number of sample data.

The primary motivation for this paper is to unwrap the mechanism causing this brittleness by providing (i) a quantitative version of the Brittleness Theorem [30, Thm. 4.13] and (ii) a detailed and comprehensive analysis of its application to the informative exam-

ple from [30, Ex. 4.16] of estimating the mean of a random variable on the unit interval using priors that exactly capture the distribution of an arbitrary large number of Hausdorff moments. In this example, the probability distribution μ^\dagger of X is an unknown element of the set of all possible probability distributions on $[0, 1]$, i.e. $\mu^\dagger \in \mathcal{A} := \mathcal{M}([0, 1])$. The set of prior probability distributions π on $\mu \in \mathcal{A}$ (i.e. $\pi \in \mathcal{M}(\mathcal{A})$) is defined as the set of priors π under which the vector of truncated Hausdorff moments $(\mathbb{E}_\mu[X], \dots, \mathbb{E}_\mu[X^n])$ is uniformly distributed on the truncated Hausdorff moment set $M^n \subset \mathbb{R}^n$ defined as the set of $q = (q_1, \dots, q_n) \in \mathbb{R}^n$ such that there exists a probability measure μ on $[0, 1]$ with $\mathbb{E}_\mu[X^i] = q_i$ for $i \in \{1, \dots, n\}$ (M^n is the polytope of \mathbb{R}^n corresponding to the set of possible values for the first n moments of a measure of probability on the interval $[0, 1]$). In this case, the computation of optimal bounds on posterior values leads naturally to the calculation of the Lebesgue volume of certain subsets of the set M^n of truncated Hausdorff moments.

Curiously, whereas Karlin and Shapley [24] used Selberg’s integral formula to first compute the volume of the truncated Hausdorff moment space M_n , inadvertently stimulating the development of the theory of the Selberg integral formulas¹, it appears that computing the volume of the truncated Hausdorff moment space M^n using different finite-dimensional representations of M^n in the infinite-dimensional space $\mathcal{M}([0, 1])$ reveals a new family of Selberg integral formulas (see Theorems 8.1, 8.2, 8.4 and Corollary 8.3). This process of discovery appears to be generic and we will now describe its main principles.

Let Ψ be the function mapping each measure $\mu \in \mathcal{M}([0, 1])$ into its first n moments

$$\Psi(\mu) := (\mathbb{E}_{X \sim \mu}[X], \mathbb{E}_{X \sim \mu}[X^2], \dots, \mathbb{E}_{X \sim \mu}[X^n]). \quad (1.1)$$

Note that

$$M^n := \Psi\left(\mathcal{M}([0, 1])\right). \quad (1.2)$$

The classical and new Selberg identities are obtained by computing the volume of M^n using different finite dimensional representations in $\mathcal{M}([0, 1])$. These finite dimensional representations are obtained by restricting Ψ to convex sums of Diracs, i.e. to measures $\mu \in \mathcal{M}([0, 1])$ of the form

$$\mu = \sum_{j=1}^N \lambda_j \delta_{t_j} \quad (1.3)$$

where $0 \leq t_1 < \dots < t_N \leq 1$ and $\lambda_1, \dots, \lambda_N > 0$ with $\sum_{j=1}^N \lambda_j = 1$. Note that if μ is of the form (1.3), then $\Psi(\mu) = (q_1, \dots, q_n)$ with $q_i = \sum_{j=1}^N \lambda_j t_j^i$.

For each measure μ of the form (1.3), we define $i(\mu)$, the index of μ , as the number of support points (Diracs) of μ , counting interior points with weight 1 and boundary points with weight 1/2. We call μ “principal” if $i(\mu) = (n+1)/2$, “canonical” if $i(\mu) = (n+2)/2$,

¹In discussing the history and importance of the Selberg integral formulas, Forrester and Waardan [19, Pg. 3] mention their first application: “For over thirty years the Selberg integral went essentially unnoticed. It was used only once-in the special case $\alpha = \beta = 1, \gamma = 2$ -in a study by S. Karlin and L.S. Shapley relating to the volume of a certain moment space, published in 1953.”

“upper” if support points include 1, “lower” if support points do not include 1. Then Theorem 5.1 asserts that each $q \in \text{Int}(M^n)$ has a unique upper and lower principal representation. Since the volume of M_n is independent of the representation used to compute it, computing that volume with a lower and an upper representation leads to an equality corresponding to classical Selberg identities.

Now let $t_* \in (0, 1)$. Theorem 5.2 asserts that every point in the interior of M^n has a unique canonical representation whose support contains t_* , and when $t_* = 0$ or 1 , then there exists a unique principal representation whose support contains t_* . Since the volume of M_n , and the higher order moments of the uniform measure restricted to M_n , are independent of the representation used to compute them, computing these “moment moments” for all possible values of t^* leads to a family of equalities corresponding to new Selberg integral formulas and Reproducing Kernel Hilbert Spaces. Consequently, the free parameter t_* associated with Markov and Kreĭn’s canonical representations of truncated Hausdorff moments (see Section 5) which, along with their principal representations, so handily provides us with the means to prove the quantitative Brittleness Theorem 3.6, is found to generate reproducing kernel identities corresponding to reproducing kernel Hilbert spaces of polynomials (see sections 6 and 7). Furthermore, these reproducing identities lead to biorthogonal systems of Selberg integral formulas described in Theorems 8.1, 8.2 and 8.4 (see also Corollary 8.3).

Moreover, although not done here, this process can easily be generalized in simple ways. For example, the argument is valid using any measure on the moment space, not just the uniform measure, and so the introduction of alternatives for which the integrals can likewise be computed, can be used. In addition, it also appears possible that this process can be repeated with multiple free parameters $t_{*,1}, \dots, t_{*,k}$ to obtain even richer classes of (new) Selberg integral formulas.

2 OUQ with Priors

To understand OUQ one simply starts with Čebyšev [27, Pg. 4] “ Given: length, weight, position of the centroid and moment of inertia of a material rod with a density varying from point to point. It is required to find the most accurate limits for the weight of a certain segment of this rod.” According to Kreĭn [27], although Čebyšev did solve this problem, it was his student Markov who supplied the proof in his thesis. See Kreĭn [27] for an account of the history of this subject along with substantial contributions by Kreĭn. We take this mindset and apply it to more complex problems, extending the base space to functions and measures, and, instead of developing sophisticated mathematical solutions, develop optimization problems and reductions, so that their solution may be implemented on a computer, as in Bertsimas and Popescu’s [8] convex optimization approach to Čebyšev inequalities, and the Decision Analysis framework of Smith [40]. In addition to the determination of optimal bounds as a function of available information and assumptions, the OUQ methodology has the substantial benefit of demanding that different components of an organization work together to come up with information and assumptions that, together, they believe in.

Let us begin with a general formulation of OUQ with priors, where the base assumptions are sets of (function, measure) pairs and the secondary assumptions are sets of priors, that is, sets of probability measures defined on the base assumption set. Later, when we apply to Bayesian inference, we will restrict to a base assumption set consisting of a set of measures and a secondary assumption consisting of a set probability measures on the base assumption set. To that end, let \mathcal{X} be a topological space, $\mathcal{M}(\mathcal{X})$ the space of Borel probability measures on \mathcal{X} , and let $\mathcal{G} \subseteq \mathcal{F}(\mathcal{X})$ be a subset of the real-valued measurable functions $\mathcal{F}(\mathcal{X})$ on \mathcal{X} . Let \mathcal{A} be an arbitrary subset of $\mathcal{G} \times \mathcal{M}(\mathcal{X})$, and let $\Phi: \mathcal{G} \times \mathcal{M}(\mathcal{X}) \rightarrow \mathbb{R}$ be a function producing a quantity of interest. In the context of uncertainty quantification one is interested in estimating $\Phi(f^\dagger, \mu^\dagger)$, where $(f^\dagger, \mu^\dagger) \in \mathcal{G} \times \mathcal{M}(\mathcal{X})$ corresponds to an *unknown reality*. If \mathcal{A} represents all that is known about (f^\dagger, μ^\dagger) (in the sense that $(f^\dagger, \mu^\dagger) \in \mathcal{A}$ and that any $(f, \mu) \in \mathcal{A}$ could, a priori, be (f^\dagger, μ^\dagger) given the available information) then [31] shows that the quantities

$$\mathcal{U}(\mathcal{A}) := \sup_{(f, \mu) \in \mathcal{A}} \Phi(f, \mu) \quad (2.1)$$

$$\mathcal{L}(\mathcal{A}) := \inf_{(f, \mu) \in \mathcal{A}} \Phi(f, \mu) \quad (2.2)$$

determine the inequality

$$\mathcal{L}(\mathcal{A}) \leq \Phi(f^\dagger, \mu^\dagger) \leq \mathcal{U}(\mathcal{A}) \quad (2.3)$$

to be optimal with respect to the available information (i.e. $(f^\dagger, \mu^\dagger) \in \mathcal{A}$) as follows: First, it is simple to see that the inequality (2.3) follows from $(f^\dagger, \mu^\dagger) \in \mathcal{A}$. Moreover, for any $\varepsilon > 0$ there exists a $(f, \mu) \in \mathcal{A}$ such that

$$\mathcal{U}(\mathcal{A}) - \varepsilon < \Phi(f, \mu) \leq \mathcal{U}(\mathcal{A}).$$

Consequently since all that we know about (f^\dagger, μ^\dagger) is that $(f^\dagger, \mu^\dagger) \in \mathcal{A}$, it follows that the upper bound $\Phi(f^\dagger, \mu^\dagger) \leq \mathcal{U}(\mathcal{A})$ is the best obtainable given that information. The lower bound is clearly optimal in the same sense.

A classical example of a quantity of interest is the validation and certification quantity $\Phi(f, \mu) := \mu[f \geq a]$ where a is a safety margin. In the certification context one is interested in showing that $\mu^\dagger[f^\dagger \geq a] \leq \epsilon$ where ϵ is a safety certification threshold (i.e. the maximum acceptable μ^\dagger -probability of the system f^\dagger exceeding the safety margin a). If $\mathcal{U}(\mathcal{A}) \leq \epsilon$, then the system associated with (f^\dagger, μ^\dagger) is safe (given the information represented by \mathcal{A}). If $\mathcal{L}(\mathcal{A}) > \epsilon$, then the system associated with (f^\dagger, μ^\dagger) is unsafe. If $\mathcal{L}(\mathcal{A}) \leq \epsilon < \mathcal{U}(\mathcal{A})$, then the safety of the system cannot be decided without making further assumptions or gathering further information.

Although the OUQ optimization problems (2.1) and (2.2) are extremely large, we have shown in [31] that an important subclass enjoys significant and practical finite-dimensional reduction properties. In particular, for assumption sets corresponding to linear inequality constraints on generalized moments, the search can be reduced to one over probability measures that are products of finite convex combinations of Dirac masses with explicit upper bounds on the number of Dirac masses.

To incorporate priors, we define a prior π to be a probability measure $\pi \in \mathcal{M}(\mathcal{A})$, and define the value $\bar{\Phi}(\pi)$ of π through the extended quantity of interest $\bar{\Phi} : \mathcal{M}(\mathcal{A}) \rightarrow \mathbb{R}$ defined by

$$\bar{\Phi}(\pi) := \mathbb{E}_\pi[\Phi], \quad \pi \in \mathcal{M}(\mathcal{A}).$$

We will defer the nontrivial and not uninteresting topics of measurability to when we analyze the full OUQ with priors framework, but note that Ressel [33] has established important and relevant results for us already, in particular the measurability of the validation and certification quantity of interest discussed above under mild conditions.

We call the value $\mathbb{E}_\pi[\Phi]$ the *prior value*, and for a family of priors $\Pi \subset \mathcal{M}(\mathcal{A})$ we note that the values

$$\mathcal{U}(\Pi) := \sup_{\pi \in \Pi} \mathbb{E}_\pi[\Phi] \tag{2.4}$$

$$\mathcal{L}(\Pi) := \inf_{\pi \in \Pi} \mathbb{E}_\pi[\Phi] \tag{2.5}$$

form a natural generalization of the notations $\mathcal{U}(\mathcal{A})$ and $\mathcal{L}(\mathcal{A})$. Moreover, in the same way that $\mathcal{U}(\mathcal{A})$ and $\mathcal{L}(\mathcal{A})$ are optimal upper and lower bounds on $\Phi(f^\dagger, \mu^\dagger)$ given the information that $(f^\dagger, \mu^\dagger) \in \mathcal{A}$, $\mathcal{U}(\Pi)$ and $\mathcal{L}(\Pi)$ are optimal upper and lower bounds on $\mathbb{E}_\pi[\Phi]$ given the information that $\pi \in \Pi$.

For conditioning on sample data in an observation space \mathcal{D} , we begin by defining a *data map*

$$\mathbb{D} : \mathcal{A} \rightarrow \mathcal{M}(\mathcal{D})$$

which specifies that $\mathbb{D}(f, \mu) \in \mathcal{M}(\mathcal{D})$ generates the data when the truth is $(f, \mu) \in \mathcal{A}$. Then, given a prior $\pi \in \mathcal{M}(\mathcal{A})$, we define a probability measure

$$\pi \odot \mathbb{D} \in \mathcal{M}(\mathcal{A} \times \mathcal{D})$$

through

$$\pi \odot \mathbb{D}[A \times B] = \mathbb{E}_{(f, \mu) \sim \pi}[\mathbb{1}_A(f, \mu) \mathbb{D}(f, \mu)[B]], \quad A \in \mathcal{B}(\mathcal{A}), B \in \mathcal{B}(\mathcal{D}),$$

where $\mathbb{1}_A$ is the indicator function of the set A :

$$\mathbb{1}_A(f, \mu) := \begin{cases} 1, & (f, \mu) \in A, \\ 0, & (f, \mu) \notin A. \end{cases}$$

We denote the resulting \mathcal{D} -marginal by $\pi \cdot \mathbb{D} \in \mathcal{M}(\mathcal{D})$ which satisfies

$$\pi \cdot \mathbb{D}[B] := \mathbb{E}_{(f, \mu) \sim \pi}[\mathbb{D}(f, \mu)[B]].$$

Given an observation $d \in \mathcal{D}$, to simultaneously avoid the ill-definedness of regular conditional probabilities and incorporate uncertainty in the observation process, we consider conditioning on an open subset B containing d such that $\pi \cdot \mathbb{D}[B] > 0$. The naturality of this positivity condition is fully discussed in [30], in particular it is easy to show that if B is an open ball of center δ around the data d (noted B_δ) and if the

data is randomized and distributed according to $\pi \cdot \mathbb{D}$, then the probability of the event $\pi \cdot \mathbb{D}[B_\delta] > 0$ is one. It is also shown in [30] that if the probability of the data is uniformly bounded, in the Bayesian model class \mathcal{A} , from above and below by that of a reference measure (e.g., for all $(f, \mu) \in \mathcal{A}$, $\frac{1}{\alpha} \mathbb{D}(f_0, \mu_0) \leq \mathbb{D}(f, \mu)[B_\delta] \leq \alpha \mathbb{D}(f_0, \mu_0)$ for some reference measure $\mathbb{D}(f_0, \mu_0)$) then learning and robustness appear as antagonistic properties: when $\alpha = 1$, the data is equiprobable under all measures in the model class, posterior values are equal to prior values, the method is robust but learning is not possible, and as α deviates from, learning becomes possible (posterior values depend on the data) but the method becomes increasingly brittle (the range of posterior values converges towards that of the quantity of interest Φ).

To simplify notation, we henceforth drop the notational dependence of the set B on the point d . The conditional expectation, given a prior π and data map \mathbb{D} , conditioned on a subset $B \in \mathcal{B}(\mathcal{D})$ such that $\pi \cdot \mathbb{D}[B] > 0$, is

$$\mathbb{E}_{\pi \odot \mathbb{D}}[\Phi|B] = \frac{\mathbb{E}_{(f, \mu) \sim \pi}[\Phi(f, \mu) \mathbb{D}(f, \mu)[B]]}{\mathbb{E}_{(f, \mu) \sim \pi}[\mathbb{D}(f, \mu)[B]]}.$$

To represent uncertainty regarding the data generating process, instead of a single data map $\mathbb{D} : \mathcal{A} \rightarrow \mathcal{M}(\mathcal{D})$, we instead specify a set

$$\mathfrak{D} = \{\mathbb{D} : \mathcal{A} \rightarrow \mathcal{M}(\mathcal{D})\}$$

of data maps and represent our assumptions regarding the data with the statement $\mathbb{D} \in \mathfrak{D}$. Therefore, having specified a set Π of priors, and a set \mathfrak{D} of data maps, for an open subset $B \subseteq \mathcal{D}$, we define the set of all possible resulting product measures to be

$$\Pi \odot_B \mathfrak{D} := \left\{ \pi \odot \mathbb{D} : \pi \in \Pi, \mathbb{D} \in \mathfrak{D}, (\pi \cdot \mathbb{D})[B] > 0 \right\}.$$

The notations $\mathcal{U}(\Pi)$ and $\mathcal{L}(\Pi)$ of (2.4) and (2.5) extend naturally to these conditional expectations as

$$\begin{aligned} \mathcal{U}(\Pi \odot_B \mathfrak{D}) &:= \sup_{\pi \odot \mathbb{D} \in \Pi \odot_B \mathfrak{D}} \mathbb{E}_{\pi \odot \mathbb{D}}[\Phi|B] \\ \mathcal{L}(\Pi \odot_B \mathfrak{D}) &:= \inf_{\pi \odot \mathbb{D} \in \Pi \odot_B \mathfrak{D}} \mathbb{E}_{\pi \odot \mathbb{D}}[\Phi|B], \end{aligned}$$

where we note that, just as for $\mathcal{U}(\mathcal{A})$ and $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\Pi)$ and $\mathcal{L}(\Pi)$, $\mathcal{U}(\Pi \odot_B \mathfrak{D})$ and $\mathcal{L}(\Pi \odot_B \mathfrak{D})$ are optimal upper and lower bounds on the posterior value $\mathbb{E}_{\pi \odot \mathbb{D}}[\Phi|B]$, given the assumptions that $\pi \in \Pi$, $\mathbb{D} \in \mathfrak{D}$, and $\pi \cdot \mathbb{D}(B) > 0$.

We are now prepared to discuss the brittleness theorems of the next section. Indeed, it is easy to see that

$$\mathcal{L}(\mathcal{A}) \leq \mathcal{L}(\Pi) \leq \mathcal{U}(\Pi) \leq \mathcal{U}(\mathcal{A}) \tag{2.6}$$

and

$$\mathcal{L}(\mathcal{A}) \leq \mathcal{L}(\Pi \odot_B \mathfrak{D}) \leq \mathcal{U}(\Pi \odot_B \mathfrak{D}) \leq \mathcal{U}(\mathcal{A}).$$

What the Brittleness Theorem 3.2 will show is that, under mild conditions, regardless of where the values $\mathcal{L}(\Pi)$ and $\mathcal{U}(\Pi)$ lie in (2.6) we have

$$\mathcal{L}(\Pi \odot_B \mathfrak{D}) \approx \mathcal{L}(\mathcal{A}) \text{ and } \mathcal{U}(\Pi \odot_B \mathfrak{D}) \approx \mathcal{U}(\mathcal{A}),$$

that is, conditioning on the observed data, one can obtain any value between $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$ for the posterior value $\mathbb{E}_{\pi \odot \mathbb{D}}[\Phi|B]$ for some admissible prior $\pi \in \Pi$ and data map $\mathbb{D} \in \mathfrak{D}$.

3 Quantification of Bayesian Brittleness

The following theorem is the Main Brittleness result of [30, Thm. 4.13]:

Theorem 3.1. *Let \mathcal{A} be a Suslin space, let \mathcal{Q} be a separable and metrizable space, and let $\Psi: \mathcal{A} \rightarrow \mathcal{Q}$ be measurable. Moreover, let $\mathfrak{Q} \subseteq \mathcal{M}(\mathcal{Q})$ be such that $\text{supp}(\mathbb{Q}) \subseteq \Psi(\mathcal{A})$ for all $\mathbb{Q} \in \mathfrak{Q}$. Suppose that, for all $\delta > 0$, there exists some $\mathbb{Q} \in \mathfrak{Q}$ such that*

$$\mathbb{E}_{q \sim \mathbb{Q}} \left[\inf_{\mu \in \Psi^{-1}(q)} \mathbb{D}(\mu)[B] \right] = 0 \quad (3.1)$$

and

$$\mathbb{P}_{q \sim \mathbb{Q}} \left[\sup_{\mu \in \Psi^{-1}(q), \mathbb{D}(\mu)[B] > 0} \Phi(\mu) > \sup_{\mu \in \mathcal{A}} \Phi(\mu) - \delta \right] > 0. \quad (3.2)$$

Then

$$\mathcal{U}(\Psi^{-1}\mathfrak{Q}|B) = \mathcal{U}(\mathcal{A}). \quad (3.3)$$

The following generalization of the Theorem 3.1 ([30, Thm. 4.13]) allows a weakening of its assumptions while approximately obtaining its conclusion. We require, as in [30], the data space \mathcal{D} to be metrizable. We select a consistent metric, and for a data point $d \in \mathcal{D}$, let $B_\delta(d)$ denote the open ball of metric radius δ about d . To keep the notation simple we omit reference to the base point d and denote this family of open balls about d by $\{B_\delta, \delta \geq 0\}$, where $B_0 = \emptyset$.

Theorem 3.2. *For a metrizable topological space \mathcal{X} , consider a topologized subset $\mathcal{G} \subset \mathcal{F}(\mathcal{X})$ and the space of probability measures $\mathcal{M}(\mathcal{X})$ equipped with the weak star topology. Let $\mathcal{A} \subset \mathcal{G} \times \mathcal{M}(\mathcal{X})$ be Suslin, \mathcal{Q} separable metrizable, and $\Psi: \mathcal{A} \rightarrow \mathcal{Q}$ Borel measurable. Moreover, let $\mathfrak{Q} \subset \mathcal{M}(\mathcal{Q})$ be such that $\text{supp } \mathbb{Q} \subset \Psi(\mathcal{A}), \mathbb{Q} \in \mathfrak{Q}$, and let $\tau \geq 0$. Suppose there exists some $\mathbb{Q} \in \mathfrak{Q}, \mathbb{D} \in \mathfrak{D}$ and a continuous monotonically increasing function $h: \mathbb{R}^+ \rightarrow \mathbb{R}$ with $h(0) = 0$ such that*

$$\mathbb{Q} \left(\left\{ q : \inf_{(f, \mu) \in \Psi^{-1}(q)} \mathbb{D}(f, \mu)[B_\delta] \leq \tau \right\} \right) \geq 1 - h(\delta), \quad \delta > 0. \quad (3.4)$$

Fix $\delta > 0$. If $\varepsilon \geq 0, \varepsilon' > 0$ and $\delta' > 0$ are three real numbers such that

$$\mathbb{Q} \left(\left\{ q : \sup_{(f, \mu) \in \Psi^{-1}(q), \mathbb{D}(f, \mu)[B_\delta] > \varepsilon} \Phi(f, \mu) > \sup_{(f, \mu) \in \mathcal{A}} \Phi(f, \mu) - \delta' \right\} \right) \geq \varepsilon' \quad (3.5)$$

and

$$h(\delta) + \tau \leq \frac{\varepsilon \delta' \varepsilon'}{\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A})}, \quad (3.6)$$

then we have

$$\mathcal{U}(\mathcal{A}) - 2\delta' \leq \mathcal{U}(\Psi^{-1}(\mathfrak{Q}) \odot_{B_\delta} \mathfrak{D}) \leq \mathcal{U}(\mathcal{A}). \quad (3.7)$$

If, for $\tau = 0$, there exists a $\delta^* > 0$ such that for all $\delta' > 0$ there exists some $\mathbb{Q} \in \mathfrak{Q}$, $\mathbb{D} \in \mathfrak{D}$ which satisfies (3.4) with a function h such that $h(\delta) = 0$, $\delta \leq \delta^*$, and which satisfies (3.5) with $\varepsilon = 0$, then we recover the conditions and the assertion of the Brittleness Theorem [30, Thm. 4.13] for B_δ , $\delta \leq \delta^*$.

Remark 3.3. The proof of Theorem 3.2 also leads to the following result. For a metrizable topological space \mathcal{X} , consider a topologized subset $\mathcal{G} \subset \mathcal{F}(\mathcal{X})$ and the space of probability measures $\mathcal{M}(\mathcal{X})$ equipped with the weak star topology. Let $\mathcal{A} \subset \mathcal{G} \times \mathcal{M}(\mathcal{X})$ be Suslin, \mathcal{Q} separable metrizable, and $\Psi : \mathcal{A} \rightarrow \mathcal{Q}$ Borel measurable. Moreover, let $\mathfrak{Q} \subset \mathcal{M}(\mathcal{Q})$ be such that $\text{supp } \mathbb{Q} \subset \Psi(\mathcal{A})$, $\mathbb{Q} \in \mathfrak{Q}$. It holds true that for $\delta > 0$

$$\mathcal{U}(\mathcal{A}) - v(\delta) \leq \mathcal{U}(\Psi^{-1}(\mathfrak{Q}) \odot_{B_\delta} \mathfrak{D}) \leq \mathcal{U}(\mathcal{A}). \quad (3.8)$$

where the function v is defined by

$$v(\delta) := 2 \inf \left\{ \delta' > 0 \left| \delta' \geq (\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A})) \inf_{\mathbb{Q} \in \mathfrak{Q}, \mathbb{D} \in \mathfrak{D}, \varepsilon > 0, \tau \geq 0} \frac{1 - \mathbb{Q} \left(\left\{ q : \inf_{(f, \mu) \in \Psi^{-1}(q)} \mathbb{D}(f, \mu)[B_\delta] \leq \tau \right\} \right) + \tau}{\varepsilon \mathbb{Q} \left(\left\{ q : \sup_{(f, \mu) \in \Psi^{-1}(q), \mathbb{D}(f, \mu)[B_\delta] > \varepsilon} \Phi(f, \mu) > \sup_{(f, \mu) \in \mathcal{A}} \Phi(f, \mu) - \delta' \right\} \right)} \right\} \quad (3.9)$$

for $\delta > 0$.

Remark 3.4. This brittleness is not a consequence of a lack of compactness of the admissible set. Indeed, in the following section, the primary space of measures $\mathcal{M}(I)$ is compact in the weak topology, as is any closed moment subset, and Theorem 3.6 describes a brittleness result.

Remark 3.5. It is true that this brittleness does not appear to be primarily due to the Bayesian methodology, but is valid more generally. See Bahadur and Savage [6] and Donoho [16] for similar results for statistical estimators, where it appears that the mechanism generating the instability is analogous to that investigated here.

3.1 Application to a revealing example

To demonstrate that the assumptions of Theorem 3.2 are mild, we now use it to extend the Brittleness result of [30, Ex. 4.16] to a simple but informative example. Here one is interested in estimating the mean of a random variable X with unknown distribution on the unit interval $I := [0, 1]$. Since our quantity of interest is $\mathbb{E}_{\mu^\dagger}[X]$, where μ^\dagger is an

unknown distribution on I , in the notations of Section 2, we have $\mathcal{X} := I$ (since X is a random variable on I), \mathcal{G} consists only of the identity function (this example does not involve unknown functions of X), $\mathcal{A} := \mathcal{M}(I)$ (the set of possible/admissible candidates for μ^\dagger is the set of all probability distributions μ on I), $\Phi(\mu) := \mathbb{E}_{t \sim \mu}[t]$ (our quantity of interest is the mean of the random variable X), $\mathcal{Q} := \mathbb{R}^n$ and the map $\Psi : \mathcal{M}(I) \rightarrow \mathbb{R}^n$ is the map to the truncated Hausdorff moments $\Psi(\mu) := (\mathbb{E}_{t \sim \mu}[t^i])_{i=1, \dots, n}$ (our set of prior distributions is defined by constraining the distribution of the first n Hausdorff moments in \mathbb{R}^n , for some fixed n). Furthermore \mathbb{Q} is the uniform Borel measure on \mathbb{R}^n restricted to the Hausdorff moment space $M^n := \Psi(\mathcal{M}(I))$ and then normalized to be a probability measure, that is $\Pi \subset \mathcal{M}(\mathcal{M}(I))$ is the set of prior distributions on $\mathcal{A} = \mathcal{M}(I)$ such that $\Psi\mu \in M^n$ is uniformly distributed on the space M^n of first n Hausdorff moments.

The Brittleness Theorem 3.2 implies (see [30, Ex. 4.16]) that if we observe (condition on) k independent samples from X , i.e. $\mathcal{D} := I^k$ and $\mathbb{D}^k \mu := \mu \otimes \dots \otimes \mu$ (k -fold tensorization) and B is the k -fold product of small enough balls centered on the data then $\mathcal{L}(\Pi \odot_B \mathcal{D}) \approx \mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\Pi \odot_B \mathcal{D}) \approx \mathcal{U}(\mathcal{A})$. In other words, although the set of prior values of $\mathbb{E}_\mu[X]$ is the single point $\{\frac{1}{2}\}$, the optimal bounds on the posterior values of $\mathbb{E}_\mu[X]$ are zero and one irrespective of the number n of constraints on marginals and the number k of observed samples if the data is observed with sufficient precision.

The following theorem provides a rigorous and quantitative statement and proof of this implication for $k = 1$. Although, for the sake of conciseness and clarity our analysis is provided in the $k = 1$ case, it generalizes to the situation where k is arbitrary. Indeed, although counterintuitive, one can show that brittleness for the single sample case is *more* difficult to obtain than for multiple samples. Since our main objective here is to unwrap and scrutinize the mechanism causing brittleness in Bayesian inference, we therefore chose to keep the presentation and our example as clear, concise, and simple as possible to illustrate the generic and pervasive nature of this brittleness.

Therefore, we will now (i) consider the case of a single data point, i.e., $k = 1$, $\mathcal{D} := I$, and $\mathbb{D}^1 \mu := \mu$ (ii) use Theorem 3.2 to provide quantitative bounds on $\mathcal{U}(\Pi \odot_B \mathcal{D})$ as a function n of the number of marginal constraints defining the set of priors (iii) scrutinize the brittleness causing mechanism through the proof of the following theorem.

Theorem 3.6. *Let $\mathcal{A} := \mathcal{M}(I)$, $\Phi(\mu) = \mathbb{E}_{t \sim \mu}[t]$, $\mathcal{D} = I$, and $\Psi : \mathcal{M}(I) \rightarrow \mathbb{R}^n$ denote the map to the truncated Hausdorff moments $\Psi(\mu) = (\mathbb{E}_{t \sim \mu}[t^i])_{i=1, \dots, n}$. Furthermore, let \mathbb{Q} denote the uniform Borel measure on \mathbb{R}^n restricted to the Hausdorff moment space $M^n := \Psi(\mathcal{M}(I))$ and then normalized to be a probability measure. Suppose that $\mathbb{Q} \in \mathfrak{Q}$ and $\mathbb{D}^1 \in \mathfrak{D}$. Then for $\delta > 0$ we have*

$$1 - 4e\left(\frac{2n\delta}{e}\right)^{\frac{1}{2n+1}} \leq \mathcal{U}(\Psi^{-1}(\mathfrak{Q}) \odot_{B_\delta} \mathfrak{D}) \leq 1 \quad (3.10)$$

Remark 3.7. Alternatively, Theorem 3.6 asserts that for positive δ, δ' satisfying

$$\delta \leq \frac{1}{4n}(\delta')^{2n+1}(2e)^{-2n}$$

we have

$$1 - 2\delta' \leq \mathcal{U}(\Psi^{-1}(\mathfrak{Q}) \odot_{B_\delta} \mathfrak{D}) \leq 1. \quad (3.11)$$

4 Volume Inequalities on the Hausdorff Moment Space

Karlin and Shapley [24, Thm. 15.2] (see also [25, Thm. 6.2]) computed the volume of the space of truncated Hausdorff moments M^n of probability measures on the unit interval to be

$$\text{Vol}(M^n) = \prod_{k=1}^n \frac{\Gamma(k)\Gamma(k)}{\Gamma(2k)}, \quad (4.1)$$

where Γ is the Gamma function. To accomplish this, they used a Markov representation of truncated moment points, as described in Kreĭn [27] (see also [25, Ch. II]), combined with the change of variables formula, followed by the evaluation of a Selberg integral.

Here we will refine their analysis to obtain volume *inequalities* on the Hausdorff moment space which are used in the application of the Brittleness Theorem 3.2 to the proof of the Brittleness Theorem 3.6. Of the two main results, it is interesting to note that the Mass Supremum Equality uses the canonical representation of moment points combined with Markov's Maximal Mass Theorem [27, Thm. 2.1] (see also [25, Thm. 4.1]) to change the ‘‘Inequality’’ to ‘‘Equality’’, whereas the Mass Infimum Inequality instead uses the principal representation, as in Karlin and Shapley's proof of the volume formula (4.1). All this terminology will be defined in the following Section 5 and comes from Karlin and Studden [25]. This section will simply state the volume inequalities that we need for Theorem 3.6.

To proceed, let us now fix terminology. Let $I := [0, 1]$, and let $\mathcal{P}(I)$ be the set of Borel measures on I and $\mathcal{M}(I) \subset \mathcal{P}(I)$ be the set of probability measures. Throughout we will assume the weak star topology for these measures. For the system of functions

$$u_i(t) := t^i, t \in I, i = 0, \dots, n$$

the Hausdorff moments of a measure $\mu \in \mathcal{P}(I)$ is defined as the vector $q \in \mathbb{R}^{n+1}$ with coordinates $q_i = \mathbb{E}_\mu[u_i] = \mathbb{E}_{t \sim \mu}[t^i]$. It is well known (see e.g. [4, Cor. 15.7]) that the map

$$\Psi : \mathcal{P}(I) \rightarrow \mathbb{R}^{n+1}$$

defined by $\Psi(\mu) := (\mathbb{E}_{t \sim \mu}[t^i], i = 0, \dots, n)$ is affine and continuous. Furthermore, let the Hausdorff moment space $\mathcal{M}^{n+1} \subset \mathbb{R}^{n+1}$ be the image $\mathcal{M}^{n+1} := \Psi\mathcal{P}(I)$ of the measures, and let M^n defined by $\mathcal{M}^{n+1} = (1, M^n)$ be moments of a probability measures omitting the zero-th moment. Equivalently, let $P_1 : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ denote the projection mapping (x_0, x_1, \dots, x_n) onto (x_1, \dots, x_n) and let $\Psi_1 =: P_1\Psi$. Then $M^n = \Psi_1\mathcal{M}(I)$. We will abuse notation by letting Ψ also denote the mapping Ψ_1 restricted to the first-to- n -th order moments of the probability measures

$$\Psi : \mathcal{M}(I) \rightarrow M^n \subset \mathbb{R}^n,$$

and, for $q \in \mathbb{R}^n$, let

$$\Psi^{-1}q := \{\mu \in \mathcal{M}(I) : \Psi\mu = q\}$$

denote its set-valued inverse.

It follows from continuity that the moment set \mathcal{M}^{n+1} is a closed convex cone and M^n is a compact convex set. Moreover, one can show that

$$\text{Int}(\mathcal{M}^{n+1}) \cap (1, \mathbb{R}^n) = (1, \text{Int}(M^n)), \quad (4.2)$$

see e.g. [34, Cor. 6.5.1], so that a point q is interior to M^n if and only if $(1, q)$ is interior to \mathcal{M}^{n+1} . Let Vol be the usual n -dimensional volume measure. Then, since M^n is convex, by [9, Lem. 1.8.1]

$$\text{Vol}(\text{Int}(M^n)) = \text{Vol}(M^n). \quad (4.3)$$

Our first result is the Mass Supremum Equality.

Lemma 4.1. *Let $t_* \in I$, $0 \leq \varepsilon \leq 1$, and consider the set $M_\varepsilon^n \subset M^n$ defined by*

$$M_\varepsilon^n := \{q \in M^n : \exists \mu \in \Psi^{-1}q : \mu(\{t_*\}) \geq \varepsilon.\}$$

Then we have

$$\text{Vol}(M_\varepsilon^n) = (1 - \varepsilon)^n \text{Vol}(M^n).$$

Remark 4.2. Lemma 4.1 is valid for any system $u_i : I \rightarrow \mathbb{R}, i = 1, \dots, n$ of moment functions which form a T-system per [25].

The second is the Mass Infimum Inequality.

Lemma 4.3. *Let $t_* \in I$, $\delta > 0$, and consider the set $M_\delta^n \subset M^n$ defined by*

$$M_\delta^n := \{q \in M^n : \exists \mu \in \Psi^{-1}q : \mu(B_\delta(t_*)) = 0.\}$$

Then we have

$$\text{Vol}(M_\delta^n) \geq (1 - \delta(2e)^{2n}) \text{Vol}(M^n).$$

The third is the Mass of First Moment Inequality.

Lemma 4.4. *Let $0 \leq \delta \leq \frac{1}{2}$. Then we have*

$$\delta^n (2e)^n \geq \frac{\text{Vol}(q \in M^n : q_1 \in [1 - \delta, 1])}{\text{Vol}(M^n)} \geq \delta^n.$$

5 Integral Geometry of the Markov-Kreĭn Representations

Here we will describe the Markov-Kreĭn representations of truncated moments and begin the development of their integral geometry. The history of this subject begins with Čebyšev and his student Markov's thesis, followed by work by Kreĭn and others, where in [27] one can find, not only an historical sketch, but substantial contributions by Kreĭn. Indeed, it is clear from Karlin and Studden [25] that this subject owes a lot to Kreĭn. Consequently, we refer to the (principal and canonical) representations that we use as Markov-Kreĭn representations. It can be argued that the appropriate name should be Čebyšev-Markov-Kreĭn representations but this name is too long and so we implicitly give credit to Čebyšev.

Now, following Karlin and Studden [25, Chapters II & IV], we describe the Markov-Kreĭn representations and determine their Jacobian determinants. We finish this section by setting up the change of variables approach, in preparation for both the proofs of the volume inequalities of Section 4 and all that follows. To wit, we define the index $i(t)$ of a strictly increasing set t of points $0 \leq t_1 < t_2 < \dots < t_N \leq 1$ by counting the interior points with weight 1 and boundary points with weight $\frac{1}{2}$. For a point $q \in M^n$ we say that a measure $\mu \in \mathcal{M}(I)$ is a representing measure for q if $\Psi(\mu) = q$ and it is a weighted sum of Dirac masses

$$\mu = \sum_{j=1}^N \lambda_j \delta_{t_j}, \quad \lambda_j > 0, j = 1, \dots, N$$

for a strictly increasing set of points $0 \leq t_1 < t_2 < \dots < t_N \leq 1$. In that case, we have the formula

$$q_i = (\Psi(\mu))_i = \sum_{j=1}^N \lambda_j t_j^i.$$

The index $i(\mu)$ of such a representing measure is defined to be the index $i(t)$ of its set of support points.

A representing measure μ is called *principal* if $i(\mu) = \frac{n+1}{2}$ and *canonical* if $i(\mu) \leq \frac{n+2}{2}$. For $q \in \text{Int}(M^n)$, [25, Thm. 2.1] asserts that $i(\mu) \geq \frac{n+1}{2}$ for any representing measure μ for q . A principal or canonical representation is called “lower” if its set of support points *does not* include the righthand endpoint 1 and “upper” if it does. The following two results will be our main tools. The first is the principal representation, see [25, Cor. 3.1].

Theorem 5.1. *Every point $q \in \text{Int}(M^n)$ has a unique upper and lower principal representation.*

The second is the canonical representation which allows the specification of a pre-determined point $t_* \in I$ in the support of the representing measure, see [25, Thm. 3.1] combined with [25, Cor. 3.2] and [25, Cor. 3.1].

Theorem 5.2. *For $t_* \in (0, 1)$, every point $q \in \text{Int}(M^n)$ has a unique canonical representation whose support contains t_* . When $t_* = 0$ or 1, there exists a unique principal representation whose support contains t_* .*

What Theorem 5.2 doesn't make clear is if the canonical representations converge to these principal representation as t_* tends to 0 or 1. They indeed do as we will see. Let us define some notation that we will use henceforward. We consider two coordinate representations of the interior of the regular unit simplex. In particular, let

$$T^N = \{(t_1, \dots, t_N) : 0 < t_1 < t_2 < \dots < t_N < 1\}$$

denote the set of strictly increasing sequences of length N in the interior to I and

$$\Lambda^N = \{(\lambda_1, \dots, \lambda_N) : \lambda_j > 0, j = 1, \dots, N, \sum_{j=1}^N \lambda_j < 1\}.$$

denote the interior to the positive orthant restricted to $\lambda \cdot \mathbb{1} < 1$. Sometimes it will be convenient to abuse this notation and shift indices so that

$$\Lambda^N = \{(\lambda_0, \dots, \lambda_{N-1}) : \lambda_j > 0, j = 0, \dots, N-1, \sum_{j=0}^{N-1} \lambda_j < 1\}.$$

We will often use the fact that I^N can be described by $N!$ copies of T^N corresponding to permuting the sequence.

We use the notation t for a vector with coordinates t_j and similarly λ for a vector with coordinates λ_j . We use the superscripts p for “principal” and c for “canonical”, the subscripts o for “odd”, e for “even”, l for “lower”, and u for “upper”. Finally, we purposefully ignore multiples of ± 1 in all our determinant calculations. With proper caution, this causes no harm since at the end of the day we take the absolute value.

5.1 Principal Representations

Theorem 5.1 asserts that each $q \in \text{Int}(M^n)$ has a unique upper and lower principal representation. We now define these representations as maps and compute their Jacobian determinants. We state these propositions without proof, since these proofs are very similar to those for the canonical representations of Propositions and 5.5 and 5.6.

First consider the odd case when $n = 2m - 1$. Then since $\frac{n+1}{2} = m$ is an integer, it follows that the support of any lower principal representation contains neither endpoint and the support of any upper principal representation contains both endpoints. Consequently, Theorem 5.1 implies that each point in $\text{Int}(M^{2m-1})$ has a unique lower principal representation of the form

$$\mu = \sum_{j=1}^m \lambda_j \delta_{t_j}, \quad \lambda_j > 0, j = 1, \dots, m, \quad \sum_{j=1}^m \lambda_j = 1 \quad (5.1)$$

where $0 < t_1 < t_2 < \dots < t_m < 1$. Therefore, consider the bijection

$$\phi_{ol}^p : \Lambda^{m-1} \times T^m \rightarrow \text{Int}(M^{2m-1})$$

defined by

$$\begin{aligned}\phi_{ol}^p(\lambda, t) &= \Psi\left(\sum_{j=1}^{m-1} \lambda_j \delta_{t_j} + (1 - \sum_{j=1}^{m-1} \lambda_j) \delta_{t_m}\right) \\ &= \left(\sum_{j=1}^{m-1} \lambda_j t_j^i + (1 - \sum_{j=1}^{m-1} \lambda_j) t_m^i\right)_{i=1}^{2m-1}.\end{aligned}\quad (5.2)$$

It also follows that each point in $Int(M^{2m-1})$ has a unique upper principal representation of the form

$$\mu = \lambda_0 \delta_0 + \sum_{j=1}^{m-1} \lambda_j \delta_{t_j} + \lambda_m \delta_1, \quad \lambda_j > 0, j = 0, \dots, m, \sum_{j=0}^m \lambda_j = 1 \quad (5.3)$$

where $0 < t_1 < t_2 < \dots < t_{m-1} < 1$. Therefore, consider the bijection

$$\phi_{ou}^p : \Lambda^m \times T^{m-1} \rightarrow Int(M^{2m-1})$$

defined by

$$\begin{aligned}\phi_{ou}^p(\lambda, t) &= \Psi\left(\lambda_0 \delta_0 + \sum_{j=1}^{m-1} \lambda_j \delta_{t_j} + (1 - \sum_{j=0}^{m-1} \lambda_j) \delta_1\right) \\ &= \left(\sum_{j=1}^{m-1} \lambda_j t_j^i + (1 - \sum_{j=0}^{m-1} \lambda_j) 1^i\right)_{i=1}^{2m-1}.\end{aligned}\quad (5.4)$$

For an increasing sequence $t_j < t_{j+1}$ let

$$\Delta(t) := \prod_{j < k} (t_k - t_j) \quad (5.5)$$

denote the Vandermonde determinant (see e.g. [23, Pg. 400]) of the matrix with entries $[t_j^i]$, $j = 1, \dots, N$, $i = 0, \dots, N - 1$. which we write as Δ_N to emphasize the dimension of t . We will also use the same formula for non-increasing sequences when we eventually take the absolute value.

Proposition 5.3. *When $n = 2m - 1$, the Jacobian determinants are*

$$\begin{aligned}|det(d\phi_{ol}^p)|(\lambda, t) &= \mathcal{J}_{ol}^p(t) \left(1 - \sum_{j=1}^{m-1} \lambda_j\right) \prod_{j=1}^{m-1} \lambda_j \\ |det(d\phi_{ou}^p)|(\lambda, t) &= \mathcal{J}_{ou}^p(t) \prod_{j=1}^{m-1} \lambda_j\end{aligned}$$

where

$$\mathcal{J}_{ol}^p(t) := \Delta_m^4(t)$$

$$\mathcal{J}_{ou}^p(t) := \prod_{j=1}^{m-1} t_j^2 (1-t_j)^2 \cdot \Delta_{m-1}^4(t).$$

Note that although each term appears to have the same multiplier $\prod_{j=1}^{m-1} \lambda_j$, in the lower case this multiplier is the full product in on Λ^{m-1} and in the upper case it is only a partial product on Λ^m , that is, it is missing the λ_0 term. Finally, let us observe the symmetries under the reflection $t \mapsto 1-t$:

$$\begin{aligned} \mathcal{J}_{ol}^p(1-t) &= \mathcal{J}_{ol}^p(t) \\ \mathcal{J}_{ou}^p(1-t) &= \mathcal{J}_{ou}^p(t). \end{aligned} \quad (5.6)$$

Now consider the even case when $n = 2m$. Since $\frac{n+1}{2} = m + \frac{1}{2}$ is an integer plus $\frac{1}{2}$ it follows that the support of any lower principal representation contains the left endpoint but not the right and any upper principal representation contains the right endpoint but not the left. Let us first consider the lower representation. Theorem 5.1 implies that every point in the interior $Int(M^{2m})$ has a unique lower principal representation of the form

$$\mu = \sum_{j=1}^m \lambda_j \delta_{t_j} + (1 - \sum_{j=1}^m \lambda_j) \delta_0, \quad \lambda_j > 0, j = 1, \dots, m, \quad \sum_{j=1}^m \lambda_j < 1$$

where $0 < t_1 < \dots < t_m < 1$. Therefore, consider the bijection

$$\phi_{el}^p : \Lambda^m \times T^m \rightarrow Int(M^{2m})$$

defined by

$$\phi_{el}^p(\lambda, t) = \Psi\left(\sum_{j=1}^m \lambda_j \delta_{t_j} + (1 - \sum_{j=1}^m \lambda_j) \delta_0\right) \quad (5.7)$$

$$= \left(\sum_{j=1}^m \lambda_j t_j^i\right)_{i=1}^{2m}. \quad (5.8)$$

On the other hand, every point in the interior $Int(M^{2m})$ has a unique upper principal representation of the form

$$\mu = \sum_{j=1}^m \lambda_j \delta_{t_j} + (1 - \sum_{j=1}^m \lambda_j) \delta_1, \quad \lambda_j > 0, j = 1, \dots, m, \quad \sum_{j=1}^m \lambda_j < 1$$

where $0 < t_1 < \dots < t_m < 1$. Therefore, consider the bijection

$$\phi_{eu}^p : \Lambda^m \times T^m \rightarrow Int(M^{2m})$$

defined by

$$\phi_{eu}^p(\lambda, t) = \Psi\left(\sum_{j=1}^m \lambda_j \delta_{t_j} + (1 - \sum_{j=1}^m \lambda_j) \delta_1\right) \quad (5.9)$$

$$= \left(\sum_{j=1}^m \lambda_j t_j^i + (1 - \sum_{j=1}^m \lambda_j)\right)_{i=1}^{2m}. \quad (5.10)$$

Proposition 5.4. *When $n = 2m$ the Jacobian determinants are*

$$|det(d\phi_{el}^p)|(\lambda, t) = \mathcal{J}_{el}^p(t) \prod_{j=1}^m \lambda_j$$

$$|det(d\phi_{eu}^p)|(\lambda, t) = \mathcal{J}_{eu}^p(t) \prod_{j=1}^m \lambda_j$$

where

$$\mathcal{J}_{el}^p(t) := \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t)$$

$$\mathcal{J}_{eu}^p(t) := \prod_{j=1}^m (1 - t_j)^2 \cdot \Delta_m^4(t)$$

Here, instead of the reflection $t \mapsto 1 - t$ leaving the lower and upper invariant as in the odd case (5.6), reflection swaps lower and upper;

$$\mathcal{J}_{el}^p(1 - t) = \mathcal{J}_{eu}^p(t). \quad (5.11)$$

5.2 Canonical Representations

Theorem 5.2 asserts that, when $t_* \in (0, 1)$, every point in $Int(M^n)$ has a unique canonical representation whose support contains t_* , and when $t_* \in \{0, 1\}$, it has a unique principal representation whose support contains t_* . Therefore, every point in $Int(M^n)$ has a unique representing measure μ such that

$$\mu = \sum_{j=1}^N \lambda_j \delta_{t_j}, \quad \lambda_j > 0, \quad j = 1, \dots, N, \quad \sum_{j=1}^N \lambda_j = 1 \quad (5.12)$$

such that the sequence $0 \leq t_1 < t_2 < \dots < t_N \leq 1$ contains t_* , where for $t_* \in (0, 1)$, the sequence has index $\frac{n+1}{2}$ or $\frac{n+2}{2}$, and when $t_* = 0$ or 1 , the index is $\frac{n+1}{2}$. Now let us remove t_* from the list and use the identity $\sum_{j=1}^N \lambda_j = 1$ to solve for the weight λ_{t_*} corresponding to t_* . Changing notation from $N \mapsto N + 1$ and relabeling the indices, we obtain that

$$\mu = \sum_{j=1}^N \lambda_j \delta_{t_j} + (1 - \sum_{j=1}^N \lambda_j) \delta_{t_*}, \quad \lambda \in \Lambda^N, \quad (5.13)$$

where the resulting sequence

$$0 \leq t_1 < t_2 < \dots < t_N \leq 1$$

does not contain t_* , and when $t_* \in (0, 1)$, the removal of this interior point implies that the resulting sequence has index $\frac{n-1}{2}$ or $\frac{n}{2}$ and when $t_* = 0$ or 1 , the removal of this boundary point implies that the resulting sequence has index $\frac{n}{2}$.

Consequently, for $t_* \in (0, 1)$, to represent $Int(M^n)$ we can split into four domains, two corresponding to the two ways of producing index $\frac{n-1}{2}$ and two corresponding the two ways of producing index $\frac{n}{2}$. When n is even one of the two index $\frac{n-1}{2}$ configurations corresponds to including $t = 0$ in the sequence and not $t = 1$ and the other corresponds to including $t = 1$ in the sequence and not $t = 0$, while one of the two index $\frac{n}{2}$ configurations corresponds to not allowing $t = 0$ or $t = 1$ and the other corresponds to including both $t = 0$ and $t = 1$. When n is odd this relationships is reversed. Similarly, when $t_* \in \{0, 1\}$, we can split into two domains corresponding to the two ways of producing index $\frac{n}{2}$.

However, we can show that the representations of index $\frac{n-1}{2}$ produce *zero* volume and so can be excluded from the integral analysis. To that end, we only need to consider the $t_* \in (0, 1)$ case. Then let us decompose the set of sequences of index \bar{N} by their endpoint configurations. That is, split such sequences into those which contain 0 but not 1, 1 but not 0, 0 and 1, and neither 0 or 1. Some of these components will be empty. On any of these endpoint specific subdomains let

$$\mathcal{I} \subset \{1, \dots, N\}$$

denote the indices of the interior points, so that in this notation we have

$$\mu = \sum_{j \in \mathcal{I}} \lambda_j \delta_{t_j} + \sum_{j \in \mathcal{I}^c} \lambda_j \delta_{t_j} + \left(1 - \sum_{j=1}^N \lambda_j\right) \delta_{t_*}, \quad \lambda \in \Lambda^N.$$

Moreover, for a sequence t , let \mathring{t} denote the sequence of interior points, and define $\mathring{T}_* := \{\mathring{t} : t \in T^N, t_j \neq t_*, j = 1, \dots, N\}$ to be the set of interior points which do not cover t_* and consider the map

$$\phi : \Lambda^N \times \mathring{T}_* \rightarrow Int(M^n)$$

defined by

$$\begin{aligned} \phi(\lambda, \mathring{t}) &= \Psi \left(\sum_{j \in \mathcal{I}} \lambda_j \delta_{t_j} + \sum_{j \in \mathcal{I}^c} \lambda_j \delta_{t_j} + \left(1 - \sum_{j=1}^N \lambda_j\right) \delta_{t_*} \right) \\ &= \left(\sum_{j \in \mathcal{I}} \lambda_j t_j^i + \sum_{j \in \mathcal{I}^c} \lambda_j t_j^i + \left(1 - \sum_{j=1}^N \lambda_j\right) t_*^i \right)_{i=1}^n \end{aligned}$$

where we note that the first sum $\sum_{j \in \mathcal{I}} \lambda_j t_j^i$ is over the interior points and the second $\sum_{j \in \mathcal{I}^c} \lambda_j t_j^i$ over the endpoints which are fixed.

The dimension of the domain $\Lambda^N \times \mathring{T}_*$ is clearly $N + |\mathcal{I}|$. However, one can easily show that

$$N + |\mathcal{I}| = 2\bar{N},$$

so that in the case $\bar{N} = \frac{n-1}{2}$, it follow that the dimension of this subdomain is $N + |\mathcal{I}| = 2\bar{N} = n - 1 < n$. Consequently, the image of this subdomain under the map ϕ has zero volume in M^n . Since the domain corresponding to index $\frac{n-1}{2}$ is a disjoint union of two

such subdomains, the assertion is proved. Moreover, the subset consisting of sequences which cover t_* also clearly has zero volume, so the constraint that the sequences not cover t_* can also be removed.

In conclusion, we can represent the volume $Vol(M^n)$ using the the representation

$$\sum_{j=1}^N \lambda_j t_j^i + (1 - \sum_{j=1}^N \lambda_j) t_*^i, \quad i = 1, \dots, n, \quad (5.14)$$

defined on two subdomains corresponding to the two ways that the sequence

$$0 \leq t_1 < t_2 < \dots < t_N \leq 1$$

can have index $\frac{n}{2}$. That is, when n is even, one subdomain corresponds to not allowing 0 or 1 and the other to including both 0 and 1. When n is odd, one subdomain corresponds to including 0 and not 1 and the other to including 1 and not 0.

We now compute the Jacobian determinants. First consider the odd case, $n = 2m - 1$. Then, sequences of index $\frac{n}{2} = m - \frac{1}{2}$ split into the lower and upper sequences

$$\begin{cases} 0 = t_1 < t_2 < \dots < t_m < 1, \\ 0 < t_1 < t_2 < \dots < t_m = 1. \end{cases}$$

Define the lower representation

$$\phi_{ol}^c : \Lambda^m \times T^{m-1} \rightarrow Int(M^{2m-1})$$

by

$$\begin{aligned} \phi_{ol}^c(\lambda, t; t_*) &= \Psi \left(\lambda_0 \delta_0 + \sum_{j=1}^{m-1} \lambda_j \delta_{t_j} + (1 - \sum_{j=0}^{m-1} \lambda_j) \delta_{t_*} \right) \\ &= \left(\sum_{j=1}^{m-1} \lambda_j t_j^i + (1 - \sum_{j=0}^{m-1} \lambda_j) t_*^i \right)_{i=1}^{2m-1} \end{aligned} \quad (5.15)$$

and the upper representation

$$\phi_{ou}^c : \Lambda^m \times T^{m-1} \rightarrow Int(M^{2m-1})$$

by

$$\begin{aligned} \phi_{ou}^c(\lambda, t; t_*) &= \Psi \left(\lambda_0 \delta_1 + \sum_{j=1}^{m-1} \lambda_j \delta_{t_j} + (1 - \sum_{j=0}^{m-1} \lambda_j) \delta_{t_*} \right) \\ &= \left(\lambda_0 + \sum_{j=1}^{m-1} \lambda_j t_j^i + (1 - \sum_{j=0}^{m-1} \lambda_j) t_*^i \right)_{i=1}^{2m-1}. \end{aligned} \quad (5.16)$$

Proposition 5.5. When $n = 2m - 1$, for $t_* \in (0, 1)$, the Jacobian determinants are

$$|\det(d\phi_{ol}^c)|(\lambda, t; t_*) = \mathcal{J}_{ol}^c(t_*, t) \prod_{j=1}^{m-1} \lambda_j$$

$$|\det(d\phi_{ou}^c)|(\lambda, t; t_*) = \mathcal{J}_{ou}^c(t_*, t) \prod_{j=1}^{m-1} \lambda_j$$

where

$$\begin{aligned} \mathcal{J}_{ol}^c(t_*, t) &:= t_* \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} t_j^2 \cdot \Delta_{m-1}^4(t) \\ \mathcal{J}_{ou}^c(t_*, t) &:= (1 - t_*) \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} (1 - t_j)^2 \cdot \Delta_{m-1}^4(t) \end{aligned}$$

Moreover,

$$\mathcal{J}_{ou}^c(t_*, t) = \mathcal{J}_{ol}^c(1 - t_*, 1 - t). \quad (5.17)$$

Now consider the even case, $n = 2m$. Then sequences of index $\frac{n}{2} = m$ split into the lower and upper sequences

$$\begin{cases} 0 < t_1 < t_2 < \cdots < t_m < 1 \\ 0 = t_1 < t_2 < \cdots < t_{m+1} = 1, \end{cases}$$

Therefore, we define the lower representation

$$\phi_{el}^c : \Lambda^m \times T^m \rightarrow \text{Int}(M^{2m})$$

by

$$\begin{aligned} \phi_{el}^c(\lambda, t; t_*) &= \Psi\left(\sum_{j=1}^m \lambda_j \delta_{t_j} + (1 - \sum_{j=1}^m \lambda_j) \delta_{t_*}\right) \\ &= \left(\sum_{j=1}^m \lambda_j t_j^i + (1 - \sum_{j=1}^m \lambda_j) t_*^i\right)_{i=1}^{2m}. \end{aligned} \quad (5.18)$$

and the upper representation

$$\phi_{eu}^c : \Lambda^{m+1} \times T^{m-1} \rightarrow \text{Int}(M^{2m})$$

by

$$\begin{aligned} \phi_{eu}^c(\lambda, t; t_*) &= \Psi\left(\lambda_0 \delta_0 + \sum_{j=1}^{m-1} \lambda_j \delta_{t_j} + \lambda_m \delta_1 + (1 - \sum_{j=0}^m \lambda_j) \delta_{t_*}\right) \\ &= \left(\sum_{j=1}^{m-1} \lambda_j t_j^i + \lambda_m + (1 - \sum_{j=0}^m \lambda_j) t_*^i\right)_{i=1}^{2m}. \end{aligned} \quad (5.19)$$

Proposition 5.6. *When $n = 2m$, for $t_* \in (0, 1)$, the Jacobian determinants are*

$$|det(d\phi_{el}^c)|(\lambda, t; t_*) = \mathcal{J}_{el}^c(t_*, t) \prod_{j=1}^m \lambda_j$$

$$|det(d\phi_{eu}^c)|(\lambda, t; t_*) = \mathcal{J}_{eu}^c(t_*, t) \prod_{j=1}^{m-1} \lambda_j,$$

where

$$\mathcal{J}_{el}^c(t_*, t) := \prod_{j=1}^m (t_j - t_*)^2 \cdot \Delta_m^4(t).$$

$$\mathcal{J}_{eu}^c(t_*, t) := t_*(1 - t_*) \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} t_j^2 (1 - t_j)^2 \cdot \Delta_{m-1}^4(t),$$

Finally, observe that if we extend the canonical representations to be defined for $t_* = 0, 1$ by continuity, we obtain the following relations between the canonical representations evaluated at the endpoints and the principal representations:

$$\begin{aligned} \mathcal{J}_{ol}^c(0, t) &\equiv 0 \\ \mathcal{J}_{ou}^c(0, t) &= \mathcal{J}_{ou}^p(t) \\ \mathcal{J}_{el}^c(0, t) &= \mathcal{J}_{el}^p(t) \\ \mathcal{J}_{eu}^c(0, t) &\equiv 0, \end{aligned} \tag{5.20}$$

$$\begin{aligned} |d\phi_{ol}^c(0, t)| &\equiv 0 \\ |d\phi_{ou}^c(0, t)| &= |d\phi_{ou}^p(t)| \\ |d\phi_{el}^c(0, t)| &= |d\phi_{el}^p(t)| \\ |d\phi_{eu}^c(0, t)| &\equiv 0, \end{aligned} \tag{5.21}$$

$$\begin{aligned} \mathcal{J}_{ol}^c(1, t) &= \mathcal{J}_{ou}^p(t) \\ \mathcal{J}_{ou}^c(1, t) &\equiv 0 \\ \mathcal{J}_{el}^c(1, t) &= \mathcal{J}_{eu}^p(t) \\ \mathcal{J}_{eu}^c(1, t) &\equiv 0, \end{aligned} \tag{5.22}$$

$$\begin{aligned} |d\phi_{ol}^c(1, t)| &= |d\phi_{ou}^p(t)| \\ |d\phi_{ou}^c(1, t)| &\equiv 0 \\ |d\phi_{el}^c(1, t)| &= |d\phi_{eu}^p(t)| \\ |d\phi_{eu}^c(1, t)| &\equiv 0. \end{aligned} \tag{5.23}$$

5.3 Change of variables integral representations

In Karlin and Shapley's [24, Thm. 15.2] proof of the Hausdorff moment volume formula (4.1), they used the lower principal representation ϕ_{ol}^p of (5.2) when n is odd and ϕ_{el}^p of (5.7) when n is even combined with the change of variables formula. To develop this method so that it can be used for the canonical representations, which are not bijections, it is convenient to proceed in some generality. To begin, consider a representation

$$\phi : W \rightarrow \text{Int}(M^n),$$

where $W \subset \mathbb{R}^n$ is open and ϕ is a continuously differentiable bijection. Then, since ϕ is injective, by the change of variables formula for injective differentiable mappings whose Jacobian determinant may vanish (see e.g. [41, Thm. 3.13] combined with Sard's Theorem [41, Thm. 3.14]), we conclude that

$$\text{Vol}(\phi(W)) = \int_W |d\phi|.$$

Moreover, since ϕ is surjective we have

$$\phi(W) = \text{Int}(M^n)$$

and from (4.3) we have

$$\text{Vol}(M^n) = \text{Vol}(\text{Int}(M^n))$$

so that we conclude

$$\text{Vol}(M^n) = \int_W |d\phi|. \quad (5.24)$$

To compute $\text{Vol}(M^n)$, Karlin and Shapley then evaluated the righthand side by determining the Jacobian determinant and then evaluating the resulting integral using a Selberg integral formula.

However, more can be done along these lines. Indeed, applying the full change of variables formula we obtain

$$\int_{\phi(W)} f = \int_W (f \circ \phi) |d\phi|$$

for any function $f : \phi(W) \rightarrow \mathbb{R}$ that is integrable over $\phi(W)$. In particular, since M^n is compact, it follows using the same reasoning that was applied above to the case $f \equiv 1$, that for any bounded measurable function $f : M^n \rightarrow \mathbb{R}$ we have

$$\int_{M^n} f = \int_W (f \circ \phi) |d\phi|. \quad (5.25)$$

We now apply this to the component functions $q \mapsto q_i, i = 1, \dots, n$ on M^n where we abuse notation and indicate them by the symbol q_i . It may be profitable to also consider

nonlinear functions such as $q \mapsto q_i^2$ but we will not do that here. Then, in this notation, $q_i \circ \phi = \phi^i$ and (5.25) becomes

$$\int_{M^n} q_i = \int_W \phi^i |d\phi| \quad (5.26)$$

That is, we have an integral representation of the mean Hausdorff moments.

However, to prove Lemma 4.1, instead of a principal representation, we use a *family* of canonical representations from Section 5.2. In this case, utilizing the conclusion at (5.14), the major difference with the previous discussion is that, instead of a single bijection, there are two continuously differentiable injections

$$\phi_k : W_k \rightarrow \text{Int}(M^n), \quad k = 1, 2$$

that are volume filling in the sense that

$$\text{Vol}(\text{Int}(M^n)) = \text{Vol}(\phi_1(W_1) \cup \phi_2(W_2))$$

and

$$\phi_1(W_1) \cap \phi_2(W_2) = \emptyset$$

and, instead of $W_k, k = 1, 2$ being open, there exists open sets $V_k \subset W_k, k = 1, 2$ such that

$$\text{Vol}(W_k) = \text{Vol}(V_k).$$

Then the analysis above can easily be repeated to conclude that

$$\int_{M^n} f = \int_{W_1} f \circ \phi_1 |d\phi_1| + \int_{W_2} f \circ \phi_2 |d\phi_2| \quad (5.27)$$

for any bounded measurable function $f : M^n \rightarrow \mathbb{R}$. In particular, we conclude

$$\int_{M^n} q_i = \int_{W_1} \phi_1^i |d\phi_1| + \int_{W_2} \phi_2^i |d\phi_2|, \quad (5.28)$$

our primary integration identity for the mean Hausdorff moments to be used in the next section.

6 Mean Hausdorff Moments using the Markov-Kreĭn Representations

We are now prepared to derive integral representations of the mean truncated Hausdorff moments with respect to the uniform measure on M^n and show that the canonical representations generate reproducing kernel identities corresponding to reproducing kernel Hilbert spaces of n -th degree polynomials. These identities are used in Section 8 to derive biorthogonal systems of Selberg integral formulas. The mean moments with respect

to many other Selberg-type densities can also be computed but to keep this presentation simple we will not do that here.

We will use Selberg's result (see e.g. [19])

$$S_n(\alpha, \beta, \gamma) = \prod_{j=0}^{n-1} \frac{\Gamma(\alpha + j\gamma)\Gamma(\beta + j\gamma)\Gamma(1 + (j+1)\gamma)}{\Gamma(\alpha + \beta + (n+j-1)\gamma)\Gamma(1 + \gamma)} \quad (6.1)$$

for the integrals

$$S_n(\alpha, \beta, \gamma) := \int_{I^n} \prod_{j=1}^n t_j^{\alpha-1} (1-t_j)^{\beta-1} |\Delta(t)|^{2\gamma} dt, \quad (6.2)$$

where $Re(\alpha) > 0, Re(\beta) > 0, Re(\gamma) > -\min(\frac{1}{n}, Re(\alpha)/(n-1), Re(\beta)/(n-1))$.

We begin with the volume calculation and then proceed to higher moments using the result of the volume calculation. The main idea of our approach is the following. Recall from Section 5.1 that the lower and upper principal representations are each bijections with $Int(M^n)$ so that the volume $Vol(M^n)$ can be computed using the change of variables result (5.24). For example, when $n = 2m - 1$, the lower principal representation ϕ_{ol}^p defined in (5.2) and the upper principal representation ϕ_{ol}^p defined in (5.4), along with the values of their Jacobian determinants from Proposition 5.3 produce two different integral representations for $Vol(M^n)$. Specifically, in the notation for Selberg's formulas (6.1) for the integrals (6.2), using the identity $\int_{\Lambda^{m-1}} (1 - \sum_{j=1}^{m-1} \lambda_j) \prod_{j=1}^{m-1} \lambda_j d\lambda = \frac{1}{(2m-1)!}$, the lower representation yields

$$\begin{aligned} Vol(M^{2m-1}) &= \int_{\Lambda^{m-1} \times T^m} |det(d\phi_{ol}^p)| \\ &= \left(\int_{\Lambda^{m-1}} (1 - \sum_{j=1}^{m-1} \lambda_j) \prod_{j=1}^{m-1} \lambda_j d\lambda \right) \int_{T^m} \mathcal{J}_{ol}^p \\ &= \frac{1}{(2m-1)!} \int_{T^m} \mathcal{J}_{ol}^p \\ &= \frac{1}{(2m-1)!m!} \int_{I^m} \mathcal{J}_{ol}^p \\ &= \frac{1}{(2m-1)!m!} \int_{I^m} \Delta_m^4(t) dt \\ &= \frac{1}{(2m-1)!m!} S_m(1, 1, 2). \end{aligned} \quad (6.3)$$

On the other hand, using the identity $\int_{\Lambda^m} \prod_{j=1}^{m-1} \lambda_j d\lambda = \frac{1}{(2m-1)!}$, the upper representa-

tion ϕ_{ou}^p yields

$$\begin{aligned}
Vol(M^{2m-1}) &= \int_{\Lambda^m \times T^{m-1}} |det(d\phi_{ou}^p)| \\
&= \left(\int_{\Lambda^m} \prod_{j=1}^{m-1} \lambda_j d\lambda \right) \int_{T^{m-1}} \mathcal{J}_{ou}^p \\
&= \frac{1}{(2m-1)!} \int_{T^{m-1}} \mathcal{J}_{ou}^p \\
&= \frac{1}{(2m-1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ou}^p \\
&= \frac{1}{(2m-1)!(m-1)!} \int_{I^{m-1}} \prod_{j=1}^{m-1} t_j^2 (1-t_j)^2 \cdot \Delta_m^4(t) dt \\
&= \frac{1}{(m-1)!(2m-1)!} S_{m-1}(3, 3, 2).
\end{aligned} \tag{6.4}$$

Combining the two results (6.3) and (6.4) we conclude the identity

$$\frac{1}{(m-1)!} S_{m-1}(3, 3, 2) = \frac{1}{m!} S_m(1, 1, 2)$$

which is confirmed through direct calculation.

In the even case, where $n = 2m$, we use the representation ϕ_{el}^p defined in (5.7) and its Jacobian determinant from Proposition 5.4, along with the identity $\int_{\Lambda^m} \prod_{j=1}^m \lambda_j d\lambda = \frac{1}{(2m)!}$, to conclude that

$$Vol(M^{2m}) = \frac{1}{(2m)!m!} \int_{I^m} \mathcal{J}_{el}^p = \frac{1}{(2m)!m!} S_m(3, 1, 2). \tag{6.5}$$

Using same identity, the upper representation ϕ_{eu}^p defined in (5.9) yields

$$Vol(M^{2m}) = \frac{1}{(2m)!m!} \int_{I^m} \mathcal{J}_{eu}^p = \frac{1}{(2m)!m!} S_m(1, 3, 2). \tag{6.6}$$

Equating the two we conclude that

$$S_m(1, 3, 2) = S_m(3, 1, 2)$$

which is well known from the symmetry of the Selberg formula in its first two arguments, and corresponds to the change of variables $t \mapsto 1 - t$. Consequently, we see how two different integral representations of the volume $Vol(M^n)$ generate identities.

However, the canonical representations form a *one parameter family* of representations of $Int(M^n)$ and the value $Vol(M^n)$ expressed in terms of the resulting one parameter family of integrals produces more interesting results. To see this, consider the odd case $n = 2m - 1$, and the volume filling pair of representations ϕ_{ol}^c and ϕ_{ou}^c defined in

(5.15) and (5.16) with Jacobian determinants evaluated in Proposition 5.5. Apply the modified change of variables formula (5.28) in terms of these two representations, and the identity $\int_{\Lambda^m} \prod_{j=1}^{m-1} \lambda_j d\lambda = \frac{1}{(2m-1)!}$, to obtain

$$\begin{aligned}
Vol(M^{2m-1}) &= \int_{\Lambda^m \times T^{m-1}} |d\phi_{ol}^c| + \int_{\Lambda^m \times T^{m-1}} |d\phi_{ou}^c| \\
&= \int_{\Lambda^m} \prod_{j=1}^{m-1} \lambda_j d\lambda \int_{T^{m-1}} \mathcal{J}_{ol}^c + \int_{\Lambda^m} \prod_{j=1}^{m-1} \lambda_j d\lambda \int_{T^{m-1}} \mathcal{J}_{ou}^c \\
&= \frac{1}{(2m-1)!} \int_{T^{m-1}} (\mathcal{J}_{ol}^c + \mathcal{J}_{ou}^c) \\
&= \frac{1}{(2m-1)!(m-1)!} \int_{I^{m-1}} (\mathcal{J}_{ol}^c + \mathcal{J}_{ou}^c).
\end{aligned}$$

Therefore, showing the parameters, we conclude that for $t_* \in (0, 1)$ we have

$$Vol(M^{2m-1}) = \frac{1}{(2m-1)!(m-1)!} \int_{I^{m-1}} (\mathcal{J}_{ol}^c(t_*, t) + \mathcal{J}_{ou}^c(t_*, t)) dt. \quad (6.7)$$

Since the identity (6.7) holds for all $t_* \in (0, 1)$ it generates integral identities. For the first, since the integrand is continuous in t_* we can set $t_* = 0$ to obtain

$$Vol(M^{2m-1}) = \frac{1}{(2m-1)!(m-1)!} \int_{I^{m-1}} (\mathcal{J}_{ol}^c(0, t) + \mathcal{J}_{ou}^c(0, t)) dt,$$

but from Propositions 5.5 we have

$$\mathcal{J}_{ol}^c(0, t) \equiv 0$$

and from (5.20)

$$\mathcal{J}_{ou}^c(0, t) = \mathcal{J}_{ou}^p(t)$$

so that we obtain

$$Vol(M^{2m-1}) = \frac{1}{(2m-1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ou}^p(t) dt,$$

which we already knew from the volume calculation using the principal representation (6.4). However, if we compute the first order differential invariant by differentiating (6.7) with respect to t_* at $t_* = 0$ we obtain the first integral formula of Theorem 8.1.

Now consider the even case $n = 2m$, and the volume filling pair of representations ϕ_{el}^c and ϕ_{eu}^c defined in (5.18) and (5.19) with Jacobian determinants evaluated in Proposition 5.6. Apply the modified change of variables formula (5.28) in terms of these two representations, and the identities $\int_{\Lambda^m} \prod_{j=1}^m \lambda_j d\lambda = \frac{1}{(2m)!}$ and $\int_{\Lambda^{m+1}} \prod_{j=1}^{m-1} \lambda_i d\lambda = \frac{1}{(2m)!}$

to obtain

$$\begin{aligned}
Vol(M^{2m}) &= \int_{\Lambda^m \times T^m} |d\phi_{el}^c| + \int_{\Lambda^{m+1} \times T^{m-1}} |d\phi_{eu}^c| \\
&= \left(\int_{\Lambda^m} \prod_{j=1}^m \lambda_j d\lambda \right) \int_{T^m} \mathcal{J}_{el}^c + \left(\int_{\Lambda^{m+1}} \prod_{j=1}^{m-1} \lambda_j d\lambda \right) \int_{T^{m-1}} \mathcal{J}_{eu}^c \\
&= \frac{1}{(2m)!} \int_{T^m} \mathcal{J}_{el}^c + \frac{1}{(2m)!} \int_{T^{m-1}} \mathcal{J}_{eu}^c \\
&= \frac{1}{(2m)!m!} \int_{I^m} \mathcal{J}_{el}^c + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c.
\end{aligned}$$

Showing the parameters, we conclude that for $t_* \in (0, 1)$ we have

$$Vol(M^{2m}) = \frac{1}{(2m)!m!} \int_{I^m} \mathcal{J}_{el}^c(t_*, t) dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt. \quad (6.8)$$

Setting $t_* = 0$ and using

$$\mathcal{J}_{eu}^c(0, t) \equiv 0$$

from Proposition 5.6 and

$$\mathcal{J}_{el}^c(0, t) = \mathcal{J}_{el}^p(t)$$

from (5.20), we obtain

$$\begin{aligned}
Vol(M^{2m}) &= \frac{1}{(2m)!m!} \int_{I^m} \mathcal{J}_{el}^c(0, t) dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(0, t) dt \\
&= \frac{1}{(2m)!m!} \int_{I^m} \mathcal{J}_{el}^p(t) dt \\
&= \frac{1}{(2m)!m!} \int_{I^m} \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t) dt \\
&= \frac{1}{(2m)!m!} S_m(3, 1, 2)
\end{aligned}$$

which we already knew from the volume calculation using the principal representation (6.5). However, if we compute the first order differential invariant by differentiating (6.8) with respect to t_* at $t_* = 0$ we obtain the second integral formula of Theorem 8.1.

We can now proceed to compute the mean of the moments with respect to the uniform measure on M^n using the volume identities (6.3), (6.4), (6.5), (6.6) from the principal representations and (6.7) and (6.8) from the canonical representations. From the identities (5.20), (5.21), (5.22), (5.23) connecting the canonical representations at the endpoints and the principal representations, it is clear that we can generate the integral formula for the mean moments corresponding to all the principal representations except ϕ_{α}^p by doing so using the canonical representations and then evaluating the result at the endpoints. Therefore, we move directly to the canonical representations. Let δ_0 denote the indicator function defined by $\delta_0(i) = 1, i = 0$ and $\delta_0(i) = 0$ otherwise. Using the

convention that $0^0 := 1$, the following proposition utilizes the volume equalities (6.7) and (6.8) to simultaneously express themselves and the moment equalities generated by the canonical representations. For a function $\phi : I \rightarrow \mathbb{R}$ we define the diagonal extension $\Sigma\phi : I^N \rightarrow \mathbb{R}$ by

$$(\Sigma\phi)(t) := \sum_{j=1}^N \phi(t_j), \quad t \in I^N.$$

For simple powers, we introduce the notation

$$\Sigma t^i := \sum_{j=1}^N t_j^i$$

for the power sum and note the important example

$$\Sigma t^{-1} := \sum_{j=1}^N t_j^{-1}$$

that will be used in the Selberg integral formulas of Theorem 8.1.

Proposition 6.1. *Let $n = 2m - 1$. Then for all $t_* \in I$ and $i = 0, 1, \dots, 2m - 1$ we have*

$$\begin{aligned} & \int_{M^{2m-1}} q_i - \frac{t_*^i}{2m} \text{Vol}(M^{2m-1}) \\ &= \frac{\delta_0(i)}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ol}^c(t_*, t) dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ou}^c(t_*, t) dt \\ &+ \frac{2}{(2m)!(m-1)!} \int_{I^{m-1}} \Sigma t^i \left(\mathcal{J}_{ol}^c(t_*, t) + \mathcal{J}_{ou}^c(t_*, t) \right) dt. \end{aligned}$$

Let $n = 2m$. Then for all $t_* \in I$ and $i = 0, 1, \dots, 2m$ we have

$$\begin{aligned} & \int_{M^{2m}} q_i - \frac{t_*^i}{2m+1} \text{Vol}(M^{2m}) \\ &= \frac{\delta_0(i) + 1}{(2m+1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt \\ &+ \frac{2}{(2m+1)!m!} \int_{I^m} \Sigma t^i \mathcal{J}_{el}^c(t_*, t) dt + \frac{2}{(2m+1)!(m-1)!} \int_{I^{m-1}} \Sigma t^i \mathcal{J}_{eu}^c(t_*, t) dt. \end{aligned}$$

The above technique of comparing two representations of the same volume to generate identities we now apply to the higher order moments with respect to the uniform measure on the moment space (the moment moments) by simply subtracting the integral representations of Proposition 6.1 evaluated at $t_* = 0$ from that with arbitrary $t_* \in I$. We now show that this procedure produces a reproducing kernel identity on the space of polynomials.

7 The Canonical Representations and Reproducing Kernel Hilbert Spaces of Polynomials

The integral representations of Proposition 6.1 show clear signs of the existence of reproducing kernel identities of the form

$$f(x) = \int K(x, y) f(y) dy, \quad f \in H, x \in X$$

since, in the odd case, the integrand on the righthand side Σt^i is integrated against a kernel $\mathcal{J}_{ol}^c(t_*, t) + \mathcal{J}_{ou}^c(t_*, t)$ and produces a multiple of t_*^i plus some terms. Reproducing kernel Hilbert spaces are Hilbert spaces of functions such that pointwise evaluation is continuous on the Hilbert space. They have remarkable properties, in particular, the reproducing kernel identities which can be thought of like an abstract Cauchy integral formula from complex analysis.

Let us present Proposition 6.1 in reproducing kernel form. To that end, define

$$\mathcal{H}(t_*, t) := \mathcal{J}_{ou}^c(0, t) - \mathcal{J}_{ol}^c(t_*, t) - \mathcal{J}_{ou}^c(t_*, t), \quad (7.1)$$

and note that from (5.20) we have $\mathcal{J}_{ou}^c(0, t) = \mathcal{J}_{ou}^p(t)$, so that

$$\mathcal{H}(t_*, t) = \mathcal{J}_{ou}^p(t) - \mathcal{J}_{ol}^c(t_*, t) - \mathcal{J}_{ou}^c(t_*, t).$$

Moreover, observe that the symmetries

$$\begin{aligned} \mathcal{J}_{ol}^p(1-t) &= \mathcal{J}_{ol}^p(t) \\ \mathcal{J}_{ou}^p(1-t) &= \mathcal{J}_{ou}^p(t) \\ \mathcal{J}_{ou}^c(t_*, t) &= \mathcal{J}_{ol}^c(1-t_*, 1-t) \end{aligned}$$

of (5.6) and (5.17) combined with $\mathcal{J}_{ol}^c(0, t) \equiv 0$ imply that

$$\begin{aligned} \mathcal{H}(0, t) &\equiv 0 \\ \mathcal{H}(1, t) &\equiv 0. \end{aligned} \quad (7.2)$$

and

$$\mathcal{H}(1-t_*, 1-t) = \mathcal{H}(t_*, t). \quad (7.3)$$

Let Π^n denote the space of n -th degree polynomials in one variable with real coefficients.

Theorem 7.1. *For all $\phi \in \Pi^{2m-1}$ we have*

$$\begin{aligned} \phi(t_*) Vol(M^{2m-1}) &= \frac{2}{(2m-1)!(m-1)!} \int_{I^{m-1}} (\Sigma \phi)(t) \mathcal{H}(t_*, t) dt \\ &+ \frac{\phi(0)}{(2m-1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ou}^c(t_*, t) dt \\ &+ \frac{\phi(1)}{(2m-1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ol}^c(t_*, t) dt \end{aligned} \quad (7.4)$$

and for $\phi \in \Pi^{2m}$ we have

$$\begin{aligned}\phi(t_*)Vol(M^{2m}) &= -\frac{2}{(2m)!m!} \int_{I^m} (\Sigma\phi)(t) (\mathcal{J}_{el}^c(t_*, t) - \mathcal{J}_{el}^c(0, t)) dt \\ &\quad - \frac{2}{(2m)!(m-1)!} \int_{I^{m-1}} (\Sigma\phi)(t) \mathcal{J}_{eu}^c(t_*, t) dt \\ &\quad + \frac{\phi(0)}{(2m)!m!} \int_{I^m} \mathcal{J}_{el}^c(t_*, t) dt - \frac{\phi(1)}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt.\end{aligned}$$

To integrate out the diagonal extension Σ , for any function $(t_*, t) \mapsto \mathcal{J}(t_*, t)$ we let

$$\bar{\mathcal{J}}(t_*, s) := \int \mathcal{J}(t_*, (s, t_2, \dots, t_N)) dt_2 \cdots dt_N$$

denote the marginalization to the first component of t . Now for any such function \mathcal{J} , which is invariant under the symmetric group acting on its second variable, we have

$$\int_{I^N} (\Sigma\phi)(t) \mathcal{J}(t_*, t) dt = N \int_I \phi(s) \bar{\mathcal{J}}(t_*, s) ds$$

so that we obtain the following corollary to Theorem 7.1. Let us define

$$\bar{\mathcal{G}}(t_*, s) := \bar{\mathcal{J}}_{el}^c(0, s) - \bar{\mathcal{J}}_{el}^c(t_*, s) - (m-1) \bar{\mathcal{J}}_{eu}^c(t_*, s)$$

and note that

$$\bar{\mathcal{G}}(0, s) \equiv 0$$

but

$$\bar{\mathcal{G}}(1, s) \neq 0.$$

Let $\Pi_0^n \subset \Pi^n$ denote the n -th degree polynomials $\phi \in \Pi^n$ which vanish on the boundary of I , that is, $\phi(0) = \phi(1) = 0$.

Corollary 7.2. *For all $\phi \in \Pi^{2m-1}$ we have*

$$\begin{aligned}\phi(t_*)Vol(M^{2m-1}) &= \frac{2}{(2m-1)!(m-2)!} \int_I \phi(s) \bar{\mathcal{H}}(t_*, s) ds \\ &\quad + \frac{\phi(0)}{(2m-1)!(m-2)!} \int_I \bar{\mathcal{J}}_{ou}^c(t_*, s) ds \\ &\quad + \frac{\phi(1)}{(2m-1)!(m-2)!} \int_I \bar{\mathcal{J}}_{ol}^c(t_*, s) ds\end{aligned}\tag{7.5}$$

and for $\phi \in \Pi^{2m}$ we have

$$\begin{aligned}\phi(t_*)Vol(M^{2m}) &= \frac{2}{(2m)!(m-1)!} \int_I \phi(s) \bar{\mathcal{G}}(t_*, s) ds \\ &\quad + \frac{\phi(0)}{(2m)!(m-1)!} \int_I \bar{\mathcal{J}}_{el}^c(t_*, s) ds - \frac{\phi(1)}{(2m)!(m-2)!} \int_I \bar{\mathcal{J}}_{eu}^c(t_*, s) ds.\end{aligned}$$

In particular, for the normalizations

$$\begin{aligned}\hat{\mathcal{H}} &:= \frac{1}{\text{Vol}(M^{2m-1})} \frac{2}{(2m-1)!(m-2)!} \bar{\mathcal{H}} \\ \hat{\mathcal{G}} &:= \frac{1}{\text{Vol}(M^{2m})} \frac{2}{(2m)!(m-1)!} \bar{\mathcal{G}}\end{aligned}$$

we have

$$\begin{aligned}\phi(t_*) &= \int_I \phi(s) \hat{\mathcal{H}}(t_*, s) ds, \quad \phi \in \Pi_0^{2m-1} \\ \phi(t_*) &= \int_I \phi(s) \hat{\mathcal{G}}(t_*, s) ds, \quad \phi \in \Pi_0^{2m}.\end{aligned}$$

Let us now restrict our attention to the odd case and let $L^2(I)$ denote the usual Lebesgue space corresponding to the uniform Borel measure on I . Then, it is well known, see e.g. Saitoh [36, Thm. 1, Pg. 21], that the integral operator

$$\phi \mapsto \int_I \phi(s) \hat{\mathcal{H}}(t_*, s) ds, \quad \phi \in L^2(I)$$

determines a reproducing kernel Hilbert space structure on its range with reproducing kernel

$$\mathcal{K}(r_1, r_2) := \int_I \hat{\mathcal{H}}(r_1, s) \hat{\mathcal{H}}(r_2, s) ds.$$

From the definition (7.1)

$$\mathcal{H}(t_*, t) := \mathcal{J}_{ou}^c(0, t) - \mathcal{J}_{ol}^c(t_*, t) - \mathcal{J}_{ou}^c(t_*, t)$$

and the definitions of \mathcal{J}_{ol}^c and \mathcal{J}_{ou}^c from Proposition 5.5, it follows that $\mathcal{H}(\cdot, t) \in \Pi_0^{2m-1}$, $t \in I^{m-1}$, and therefore it follows from (7.2) that $\mathcal{H}(\cdot, t) \in \Pi_0^{2m-1}$, $t \in I^{m-1}$. Consequently, by marginalization to $\hat{\mathcal{H}}$ and scalar normalization, we have

$$\hat{\mathcal{H}}(\cdot, s) \in \Pi_0^{2m-1}, \quad s \in I. \quad (7.6)$$

Therefore the range of this integral operator is contained in Π_0^{2m-1} . However, it follows from Corollary 7.2 that the range is identically Π_0^{2m-1} . Therefore we conclude that Π_0^{2m-1} is a reproducing kernel Hilbert space with kernel \mathcal{K} .

Because of Corollary 7.2, one might be tempted to think that this reproducing kernel Hilbert space structure corresponds to that which Π_0^{2m-1} inherits as the subspace $\Pi_0^{2m-1} \subset L^2(I)$, but this is not the case. Indeed, let P_{2m-1} denote the $L^2(I)$ orthogonal projection $P_{2m-1} : L^2(I) \rightarrow \Pi_0^{2m-1}$ and consider the kernel

$$\mathcal{K}_{2m-1}(r_1, r_2) := \int_I P_{2m-1} \hat{\mathcal{H}}(r_1, s) \cdot P_{2m-1} \hat{\mathcal{H}}(r_2, s) ds \quad (7.7)$$

where the projections are acting on the kernels in the second component. Then, since this projection makes no difference in the reproducing identities in Corollary 7.2, one

can show that \mathcal{K}_{2m-1} is the reproducing kernel associated with $\Pi_0^{2m-1} \subset L^2(I)$ and since the latter can be computed in terms of the Legendre polynomials of order 2 (see e.g. [5, Sec. 12.5]) using the Christoffel-Darboux formula [12] (see e.g. Simon [38] for a more current reference), we conclude an identification of \mathcal{K}_{2m-1} with the Christoffel-Darboux formula for the kernel of the Legendre polynomials of order 2. That, is

$$\mathcal{K}_{2m-1}(r_1, r_2) = \frac{(r_1 - r_1^2)(r_2 - r_2^2)}{2(2m-1)(2m)(2m+1)} \frac{P_{2m}''(r_1)P_{2m-1}''(r_2) - P_{2m-1}''(r_1)P_{2m}''(r_2)}{r_1 - r_2} \quad (7.8)$$

for $(r_1, r_2) \in I^2$, where P_k are the Legendre polynomials shifted to the interval

$$P_k(r) = \frac{1}{k!} \frac{d^k(r^2 - r)^k}{dr^k}, \quad r \in I \quad (7.9)$$

and

$$Q_k(r) := (r - r^2)P_k''(r), \quad r \in I \quad (7.10)$$

are the associated Legendre polynomials of order 2 (see e.g. [5, Sec. 12.5]).

Moreover, since

$$\bar{\mathcal{H}}(t_*, s) := \bar{\mathcal{J}}_{ou}^c(0, s) - \bar{\mathcal{J}}_{ol}^c(t_*, s) - \bar{\mathcal{J}}_{ou}^c(t_*, s),$$

and from Proposition 5.5 we have

$$\begin{aligned} \bar{\mathcal{J}}_{ou}^c(0, 0) &\equiv 0 \\ \bar{\mathcal{J}}_{ol}^c(t_*, 0) &\equiv 0 \\ \bar{\mathcal{J}}_{ou}^c(t_*, 0) &> 0, \quad t_* \in (0, 1), \end{aligned}$$

we find that

$$\bar{\mathcal{H}}(t_*, 0) = -\bar{\mathcal{J}}_{ou}^c(t_*, 0) < 0, \quad t_* \in (0, 1).$$

Consequently, for $t_* \in (0, 1)$, it follows that $\hat{\mathcal{H}}(t_*, \cdot) \notin \Pi_0^{2m-1}$ and therefore

$$\mathcal{K} \neq \mathcal{K}_{2m-1}.$$

Moreover, from the orthogonal decomposition

$$\begin{aligned} \mathcal{K}(r_1, r_2) &= \int_I \hat{\mathcal{H}}(r_1, s) \hat{\mathcal{H}}(r_2, s) ds \\ &= \int_I P_{2m-1} \hat{\mathcal{H}}(r_1, s) \cdot P_{2m-1} \hat{\mathcal{H}}(r_2, s) ds + \int_I P_{2m-1}^\perp \hat{\mathcal{H}}(r_1, s) \cdot P_{2m-1}^\perp \hat{\mathcal{H}}(r_2, s) ds \\ &= \mathcal{K}_{2m-1}(r_1, r_2) + \int_I P_{2m-1}^\perp \hat{\mathcal{H}} \cdot P_{2m-1}^\perp \hat{\mathcal{H}}(r_2, s) ds \end{aligned}$$

we conclude

Theorem 7.3. Let \mathcal{K}_{2m-1} denote the reproducing kernel for the polynomials Π_0^{2m-1} as a subset of $L^2(I)$. Then \mathcal{K}_{2m-1} can be expressed by both (7.7) and (7.8). Moreover, Π_0^{2m-1} is also a reproducing kernel Hilbert space with kernel \mathcal{K} , and

$$\mathcal{K} - \mathcal{K}_{2m-1}$$

is a reproducing kernel.

Remark 7.4. Besides the fact that the kernel \mathcal{K} defining the Hilbert space structure for the polynomials Π_0^{2m-1} is not that of the Legendre polynomials, we do not know if this kernel is known, nor do we have an explicit formula for it. However, what this section shows is that this kernel and its associated Hilbert space Π_0^{2m-1} are intimately connected with the canonical representations of truncated Hausdorff moments, and therefore might be called the Markov-Kreĭn kernel. Moreover, if instead of the uniform measure on the moments, a Selberg type density is used, more such reproducing kernels may be revealed.

8 New Selberg Integral Formulas

The integral representations of the mean Hausdorff moments of Proposition 6.1 provide new integral identities of Selberg type. In the following theorem, we provide the first in a sequence corresponding to when n is odd and even. We then show how to use the reproducing kernel identities of Theorem 7.1 to generate biorthogonal systems of Selberg integral formulas.

Theorem 8.1. It holds true that

$$\int_{I^m} \Sigma t^{-1} \cdot \prod_{j=1}^m t_j^2 (1 - t_j)^2 \Delta_m^4(t) dt = \frac{S_m(5, 1, 2) - S_m(3, 3, 2)}{2}. \quad (8.1)$$

and

$$\int_{I^m} \Sigma t^{-1} \cdot \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t) dt = \frac{m}{2} S_{m-1}(5, 3, 2). \quad (8.2)$$

The identities of Theorem 8.1 follow only from the volume equalities, that is, the $i = 0$ case of Theorem 7.1. The following theorem demonstrates how to use all the moment equalities of Theorem 7.1 to generate biorthogonal systems of Selberg integral formulas. Let us recall definition (7.1)

$$\mathcal{H}(t_*, t) = \mathcal{J}_{ou}^p(t) - \mathcal{J}_{ol}^c(t_*, t) - \mathcal{J}_{ou}^c(t_*, t).$$

Theorem 8.2. Let $n = 2m - 1$ and consider the scaled kernel

$$\hat{\mathcal{H}} := \frac{1}{\text{Vol}(M^{2m-1})} \frac{2}{(2m-1)!(m-1)!} \mathcal{H}.$$

Then,

$$\hat{\mathcal{H}}(\cdot, t) \in \Pi_0^{2m-1}, \quad t \in I^{m-1}.$$

Moreover, consider a basis $\{p_j, j = 1, \dots, 2m - 2\}$ for Π_0^{2m-1} and the resulting expansion of $\hat{\mathcal{H}}(\cdot, t)$ in this basis for each $t \in I^{m-1}$;

$$\hat{\mathcal{H}}(t_*, t) = \sum_{j=1}^{2m-2} h_j(t) p_j(t_*), \quad (t_*, t) \in I \times I^{m-1}.$$

Then, $\{\Sigma p_j, j = 1, \dots, 2m - 2\}$, $\{h_j, j = 1, \dots, 2m - 2\}$ form an $L^2(I^{m-1})$ biorthogonal system. That is,

$$\int_{I^{m-1}} h_j \Sigma p_k = \delta_{jk}, \quad j, k = 1, \dots, 2m - 2$$

As an immediate corollary, we have

Corollary 8.3. *Let $n = 2m - 1$ and consider the scaled marginal kernel*

$$\hat{\mathcal{H}} := \frac{1}{\text{Vol}(M^{2m-1})} \frac{2}{(2m-1)!(m-2)!} \bar{\mathcal{H}}$$

(note the different scaling than Theorem 8.2). Then,

$$\hat{\mathcal{H}}(\cdot, s) \in \Pi_0^{2m-1}, \quad s \in I.$$

Moreover, consider a basis $\{p_j, j = 1, \dots, 2m - 2\}$ for Π_0^{2m-1} and the resulting expansion of $\hat{\mathcal{H}}(\cdot, s)$ in this basis for each $s \in I$;

$$\hat{\mathcal{H}}(t_*, s) = \sum_{j=1}^{2m-2} \bar{h}_j(s) p_j(t_*), \quad (t_*, s) \in I \times I.$$

Then, $\{p_j, j = 1, \dots, 2m - 2\}$, $\{\bar{h}_j, j = 1, \dots, 2m - 2\}$ form an $L^2(I)$ biorthogonal system. That is,

$$\int_I \bar{h}_j p_k = \delta_{jk}, \quad j, k = 1, \dots, 2m - 2.$$

The choice of basis for Π_0^{2m-1} determines the corresponding component functions $h_j, j = 1, \dots, 2m - 2$ and the integrands $h_j \Sigma p_k$ in Theorem 8.2. Therefore, the task remaining is to select a basis for which the component functions h_j can be determined and such that the resulting integrals are of interest. When the chosen basis is orthonormal with respect to some inner product $\langle \cdot, \cdot \rangle$, then the coefficients h_j in the representation

$$\hat{\mathcal{H}}(t_*, t) = \sum_{j=1}^{2m-2} h_j(t) p_j(t_*), \quad (t_*, t) \in I \times I^{m-1}.$$

of Theorem 8.1 are

$$h_j(t) = \langle \hat{\mathcal{H}}(\cdot, t), p_j \rangle.$$

As an example, we now compute these component functions, and therefore determine explicit forms for these Selberg integrals, when the basis consists of the associated

Legendre polynomials of order 2. To that end, recall the definitions (7.9) and (7.10) of the Legendre polynomials and the associated Legendre polynomials of order 2 translated to the unit interval I . In addition, recall the j -th symmetric function e_j defined as

$$e_j(t) := \sum_{i_1 < \dots < i_j} t_{i_1} \cdots t_{i_j}$$

with $e_0 := 1$ and the symmetric functions $e_j(t, z)$ restricted to the diagonal $t = z$

$$e_j(t, t) := \sum_{j_1 + j_2 = j} e_{j_1}(t) e_{j_2}(t), \quad j = 0, \dots, 2m - 2. \quad (8.3)$$

Theorem 8.4. *Consider the basis of Π_0^{2m-1} consisting of the associated Legendre polynomials $Q_j, j = 2, \dots, 2m - 1$ of order 2 translated to the unit interval I . For $k = 2, \dots, 2m - 1$ define*

$$a_{jk} := \frac{(j + k + k^2)\Gamma(j + 2)\Gamma(j)}{\Gamma(j + k + 2)\Gamma(j - k + 1)}, \quad k \leq j \leq 2m - 1$$

$$\tilde{h}_k(t) := \sum_{j=k}^{2m-1} (-1)^{j+1} a_{jk} e_{2m-1-j}(t, t).$$

Then for $j = k \bmod 2, j, k = 2, \dots, 2m - 1$, we have

$$\int_{I^{m-1}} \tilde{h}_k(t) \Sigma Q_j(t) \prod_{j'=1}^{m-1} t_{j'}^2 \cdot \Delta_{m-1}^4(t) dt = \text{Vol}(M^{2m-1}) (2m-1)! (m-1)! \frac{(k+2)!}{(8k+4)(k-2)!} \delta_{jk}.$$

9 Proofs

9.1 Proof of Theorem 3.2

We seek to apply the nested reduction theorem [30, Thm. 4.11]. The assertion is trivially true when $\mathcal{L}(\mathcal{A}) = \mathcal{U}(\mathcal{A})$ so we can assume $\mathcal{L}(\mathcal{A}) < \mathcal{U}(\mathcal{A})$. Let us first establish that the assumptions of the theorem are well defined. To that end, note that [30, Lem. 3.10] (which follows from Castaing and Valadier [10, Lemma III.39 p. 86], which in turn follows from Saint-Beuve's [35] extension of Aumann's Selection Theorem to Suslin spaces) implies that $q \rightarrow \inf_{(f, \mu) \in \Psi^{-1}(q)} \mathbb{D}(f, \mu)[B]$ is universally measurable and hence the conditions of the theorem are well defined if we extend the definitions in the usual way when operating on universally measurable sets and functions. Similarly, since for any λ the function $(f, \mu) \mapsto (\Phi(f, \mu) - \lambda) \mathbb{D}(f, \mu)[B]$ is measurable, the function $\theta : \mathcal{Q} \rightarrow \mathbb{R}$ defined by

$$\theta(q) := \sup_{(f, \mu) \in \Psi^{-1}(q)} (\Phi(f, \mu) - \lambda) \mathbb{D}(f, \mu)[B_\delta]$$

is universally measurable.

For the proof of the theorem, let $\mathbb{Q} \in \mathfrak{Q}$, $\mathbb{D} \in \mathfrak{D}$ satisfy the assumptions, and define $\lambda := \mathcal{U}(\mathcal{A}) - 2\delta'$. Consider the events

$$\mathcal{Q}_\delta := \left\{ q : \inf_{(f,\mu) \in \Psi^{-1}(q)} \mathbb{D}(f,\mu)[B_\delta] \leq \tau \right\}$$

$$U_{\varepsilon,\delta} = \left\{ q : \sup_{(f,\mu) \in \Psi^{-1}(q), \mathbb{D}(f,\mu)[B_\delta] > \varepsilon} \Phi(f,\mu) > \sup_{(f,\mu) \in \mathcal{A}} \Phi(f,\mu) - \delta' \right\}$$

where the assumptions (3.4) and (3.5) become

$$\mathbb{Q}(\mathcal{Q}_\delta) \geq 1 - h(\delta), \quad \delta > 0$$

$$\mathbb{Q}(U_{\varepsilon,\delta}) \geq \varepsilon', \quad \delta > 0$$

Let us denote $\tau' := \tau(\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A}))$. It is easy to see that

$$\{\theta \geq -\tau'\} \supset \mathcal{Q}_\delta$$

and

$$\{\theta > \varepsilon\delta'\} \supset U_{\varepsilon,\delta}$$

and therefore

$$\begin{aligned} \mathbb{Q}(\{\theta \geq -\tau'\}) &\geq \mathbb{Q}(\mathcal{Q}_\delta) \\ &\geq 1 - h(\delta) \end{aligned}$$

and

$$\begin{aligned} \mathbb{Q}(\{\theta > \varepsilon\delta'\}) &\geq \mathbb{Q}(U_{\varepsilon,\delta}) \\ &\geq \varepsilon'. \end{aligned}$$

Since $\Phi(f,\mu) \geq \mathcal{L}(\mathcal{A})$, $(f,\mu) \in \mathcal{A}$, it follows that $|\theta| \leq \mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A})$, and so we obtain

$$\begin{aligned} \int \theta d\mathbb{Q} &= \int_{\{\theta > \varepsilon\delta'\}} \theta d\mathbb{Q} + \int_{\{\theta \leq \varepsilon\delta'\}} \theta d\mathbb{Q} \\ &> \varepsilon\delta' \mathbb{Q}(\{\theta > \varepsilon\delta'\}) + \int_{\{\theta \leq \varepsilon\delta'\}} \theta d\mathbb{Q} \\ &\geq \varepsilon\delta' \mathbb{Q}(\{\theta > \varepsilon\delta'\}) + \int_{\{\theta \leq 0\}} \theta d\mathbb{Q} \\ &\geq \varepsilon\delta' \mathbb{Q}(\{\theta > \varepsilon\delta'\}) + \int_{\{\theta < -\tau'\}} \theta d\mathbb{Q} + \int_{\{-\tau' \leq \theta \leq 0\}} \theta d\mathbb{Q} \\ &\geq \varepsilon\delta' \mathbb{Q}(\{\theta > \varepsilon\delta'\}) - (\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A})) \mathbb{Q}(\{\theta < -\tau'\}) - \tau' \mathbb{Q}(\{-\tau' \leq \theta \leq 0\}) \\ &\geq \varepsilon\delta'\varepsilon' - (\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A}))h(\delta) - \tau'. \end{aligned}$$

Therefore, for any strictly positive solution $\delta > 0$ to

$$h(\delta) + \tau \leq \frac{\varepsilon\delta'\varepsilon'}{\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A})}$$

we have

$$\mathbb{E}_{\mathbb{Q}}[\theta] = \int \theta d\mathbb{Q} > 0,$$

where we recall that the function θ depends on δ , and therefore trivially

$$\sup_{\mathbb{Q} \in \Omega, \mathbb{D} \in \mathfrak{D}} \mathbb{E}_{q \sim \mathbb{Q}} \left[\sup_{(f, \mu) \in \Psi^{-1}(q)} (\Phi(f, \mu) - \lambda) \mathbb{D}(f, \mu)[B_\delta] \right] > 0.$$

The assertion then follows from [30, Thm. 4.11].

9.2 Proof of Theorem 3.6

We will apply the Mass Supremum Equality 4.1, the Mass Infimum Inequality 4.3, and the Mass of First Moment Inequality 4.4. To that end, define the events

$$S_{\varepsilon, \delta} = \left\{ q \in M^n : \exists \mu \in \Psi^{-1}q : \mu(B_\delta) > \varepsilon \right\}$$

$$I_\delta = \left\{ q \in M^n : \exists \mu \in \Psi^{-1}q : \mu(B_\delta) = 0 \right\}$$

$$FM_{\delta'} = \left\{ q \in M^n : q_1 \in (1 - \delta', 1] \right\}$$

First observe that some endpoint conditions have zero mass. For example,

$$\mathbb{Q}(S_{\varepsilon, \delta}) := \mathbb{Q}(\{q : \exists \mu \in \Psi^{-1}q : \mu(B_\delta) > \varepsilon\}) = \mathbb{Q}(\{q : \exists \mu \in \Psi^{-1}q : \mu(B_\delta) \geq \varepsilon\})$$

and

$$\mathbb{Q}(FM_{\delta'}) := \mathbb{Q}(\{q \in M^n : q_1 \in (1 - \delta', 1]\}) = \mathbb{Q}(\{q \in M^n : q_1 \in [1 - \delta', 1]\}).$$

Consequently, the Mass Supremum Equality 4.1 asserts that

$$\mathbb{Q}(S_{\varepsilon, \delta}) = \mathbb{Q}(\{q : \exists \mu \in \Psi^{-1}q : \mu(B_\delta) \geq \varepsilon\}) \geq (1 - \varepsilon)^n$$

where the right-hand side is independent of δ , the Mass Infimum Inequality 4.3 asserts that

$$\mathbb{Q}(I_\delta) \geq 1 - \delta(2e)^{2n},$$

and the Mass of First Moment Inequality 4.4 asserts that

$$\mathbb{Q}(FM_{\delta'}) = \mathbb{Q}(\{q \in M^n : q_1 \in [1 - \delta', 1]\}) \geq (\delta')^n.$$

Define the events

$$\mathcal{Q}_\delta := \left\{ q : \inf_{\mu \in \Psi^{-1}(q)} \mu[B_\delta] = 0 \right\}$$

$$U_{\varepsilon, \delta} := \left\{ q : \sup_{\mu \in \Psi^{-1}(q), \mu[B_\delta] > \varepsilon} \mathbb{E}_\mu[X] > 1 - \delta' \right\}$$

Then since

$$\mathcal{Q}_\delta \supset I_\delta$$

we have

$$\mathbb{Q}(\mathcal{Q}_\delta) \geq \mathbb{Q}(I_\delta) \geq 1 - \delta(2e)^{2n}$$

and since

$$\begin{aligned} U_{\varepsilon, \delta} &= \left\{ q : \sup_{\mu \in \Psi^{-1}(q), \mu[B_\delta] > \varepsilon} \mathbb{E}_\mu[X] > 1 - \delta' \right\} \\ &= \left\{ q : q_1 \in (1 - \delta', 1], \exists \mu \in \Psi^{-1}q : \mu[B_\delta] > \varepsilon \right\} \\ &= S_{\varepsilon, \delta} \cap FM_{\delta'} \end{aligned}$$

we have

$$\begin{aligned} \mathbb{Q}(U_{\varepsilon, \delta}) &= \mathbb{Q}(S_{\varepsilon, \delta} \cap FM_{\delta'}) \\ &= 1 - \mathbb{Q}((S_{\varepsilon, \delta} \cap FM_{\delta'})^c) \\ &= 1 - \mathbb{Q}(S_{\varepsilon, \delta}^c \cup FM_{\delta'}^c) \\ &= 1 - \mathbb{Q}(S_{\varepsilon, \delta}^c) - \mathbb{Q}(FM_{\delta'}^c) \\ &= \mathbb{Q}(S_{\varepsilon, \delta}) + \mathbb{Q}(FM_{\delta'}) - 1 \\ &= (1 - \varepsilon)^n - 1 + (\delta')^n \\ &= (\delta')^n - n\varepsilon \end{aligned}$$

Consequently, if we choose $\varepsilon := \frac{(\delta')^n}{2n}$, then

$$\mathbb{Q}(U_{\varepsilon, \delta}) \geq \frac{(\delta')^n}{2},$$

so that the assumptions (3.4) and (3.5), expressed as

$$\mathbb{Q}(\mathcal{Q}_\delta) \geq 1 - h(\delta), \quad \delta > 0$$

$$\mathbb{Q}(U_{\varepsilon, \delta}) \geq \varepsilon', \quad \delta > 0,$$

are satisfied with $\varepsilon' := \frac{(\delta')^n}{2}$, $\varepsilon := \frac{(\delta')^n}{2n}$, and $h(\delta) := \delta(2e)^{2n}$. We can solve

$$\begin{aligned} h(\delta) &\leq \frac{\varepsilon \delta' \varepsilon'}{\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A})} \\ &= \frac{\varepsilon \delta' \varepsilon'}{(\delta')^n \delta' \frac{(\delta')^n}{2}} \\ &= \frac{(\delta')^{2n+1}}{4n} \end{aligned}$$

by choosing $\delta \leq \frac{1}{4n} (\delta')^{2n+1} (2e)^{-2n}$.

9.3 Proof of Lemma 4.1

For $M_\varepsilon^n := \{q \in M^n : \exists \mu \in \Psi^{-1}q : \mu(\{t_*\}) \geq \varepsilon\}$, it follows that $\Psi\mu \in M_\varepsilon^n$ if and only if $\mu = \varepsilon\delta_{t_*} + (1-\varepsilon)\mu_*$ with $\mu_* \in \mathcal{M}(I)$. For such a μ it follows that $\Psi\mu = \varepsilon\Psi\delta_{t_*} + (1-\varepsilon)\Psi\mu_*$ and therefore

$$M_\varepsilon^n = \varepsilon\Psi\delta_{t_*} + (1-\varepsilon)M^n$$

from which we conclude that

$$\text{Vol}(M_\varepsilon^n) = (1-\varepsilon)^n \text{Vol}(M^n),$$

establishing the assertion.

9.4 Proof of Lemma 4.3

First consider the odd case, $n = 2m-1$. We utilize the bijective principal representation $\phi_{ol}^p : \Lambda^{m-1} \times T^m \rightarrow \text{Int}(M^{2m-1})$ defined in (5.2) and

$$|\det(d\phi_{ol}^p)|(\lambda, t) = \mathcal{J}_{ol}^p(t) \left(1 - \sum_{j=1}^{m-1} \lambda_j\right) \prod_{j=1}^{m-1} \lambda_j$$

where

$$\mathcal{J}_{ol}^p(t) := \Delta_m^4(t)$$

from Proposition 5.3 along with the change of variables formula (5.25).

Fix $t_* \in (0, 1)$ and let

$$T_\delta^m := \{(t_1, \dots, t_m) \in T^m : t_j \notin B_\delta(t_*), j = 1, \dots, m\}$$

denote those sequences which have no point a distance less than δ from t_* . It follows that

$$M_\delta^{2m-1} \supset \phi(\Lambda^{m-1} \times T_\delta^m)$$

and therefore

$$\text{Vol}(M_\delta^{2m-1}) \geq \text{Vol}(\phi_{ol}^p(\Lambda^{m-1} \times T_\delta^m)). \quad (9.1)$$

We bound the righthand side from below using the change of variables formula (5.25) as

$$\begin{aligned} \text{Vol}(\phi_{ol}^p(\Lambda^{m-1} \times T_\delta^m)) &= \int_{\Lambda^{m-1} \times T_\delta^m} |\det(d\phi_{ol}^p)| \\ &= \int_{\Lambda^{m-1}} \left(1 - \sum_{j=1}^{m-1} \lambda_j\right) \left(\prod_{j=1}^{m-1} \lambda_j\right) d\lambda \int_{T_\delta^m} \mathcal{J}_{ol}^p, \end{aligned}$$

and then bounding

$$\begin{aligned} \int_{T_\delta^m} \mathcal{J}_{ol}^p dt &= \int_{T_\delta^m} \Delta_m^4(t) dt \\ &= \frac{1}{m!} \int_{I_\delta^m} \Delta_m^4(t) dt \end{aligned}$$

where

$$I_\delta^m := \{(t_1, \dots, t_m) \in I^m : t_j \notin B_\delta(t_*), j = 1, \dots, m\}.$$

To bound this from below we bound the integral over $(I_\delta^m)^c$ from above. To that end, let

$$I_{\delta,j}^m := \{(t_1, \dots, t_m) \in I^m : t_j \in B_\delta(t_*)\}, \quad j = 1, \dots, m,$$

so that

$$(I_\delta^m)^c = \cup_j I_{\delta,j}^m.$$

Therefore, using a union bound and the symmetry of Δ we have

$$\begin{aligned} \int_{(I_\delta^m)^c} \Delta_m^4(t) dt &= \int_{\cup_{j'=1}^m I_{\delta,j'}^m} \Delta_m^4(t) dt \\ &\leq \sum_{j'=1}^m \int_{I_{\delta,j'}^m} \Delta_m^4(t) dt \\ &= m \int_{I_{\delta,1}^m} \Delta_m^4(t) dt \\ &= m \int_{I_{\delta,1}^m} \prod_{1 \leq j < k \leq m} (t_k - t_j)^4 dt_1 \cdots dt_m \\ &\leq m \int_{I_{\delta,1}^m} \prod_{2 \leq j < k \leq m} (t_k - t_j)^4 dt_1 \cdots dt_m \\ &= m \text{Vol}(B_\delta) \int_{I^{m-1}} \prod_{2 \leq j < k \leq m} (t_k - t_j)^4 dt_2 \cdots dt_m \\ &= m \text{Vol}(B_\delta) \int_{I^{m-1}} \Delta_{m-1}^4(t) dt \\ &\leq 2m\delta S_{m-1}(1, 1, 2) \end{aligned}$$

and so obtain

$$\begin{aligned} \frac{\text{Vol}(\phi(\Lambda^{m-1} \times T_\delta^m))}{\text{Vol}(\phi(\Lambda^{m-1} \times T^m))} &= \frac{\int_{I_\delta^m} \Delta_m^4(t) dt}{\int_{I^m} \Delta_m^4(t) dt} \\ &\geq 1 - 2m\delta \frac{S_{m-1}(1, 1, 2)}{S_m(1, 1, 2)}. \end{aligned}$$

Using Selberg's formulas (6.1) we compute

$$\begin{aligned}
\frac{S_{m-1}(1, 1, 2)}{S_m(1, 1, 2)} &= \frac{\prod_{j=0}^{m-2} \frac{\Gamma(1+2j)^2 \Gamma(3+2j)}{2\Gamma(2(m+j)-2)}}{\prod_{j=0}^{m-1} \frac{\Gamma(1+2j)^2 \Gamma(3+2j)}{2\Gamma(2(m+j))}} \\
&= \frac{2\Gamma(4m-4)}{\Gamma(2m-1)^2 \Gamma(2m+1)} \frac{\prod_{j=0}^{m-1} \frac{\Gamma(1+2j)^2 \Gamma(3+2j)}{2\Gamma(2(m+j)-2)}}{\prod_{j=0}^{m-1} \frac{\Gamma(1+2j)^2 \Gamma(3+2j)}{2\Gamma(2(m+j))}} \\
&= \frac{2\Gamma(4m-4)}{\Gamma(2m-1)^2 \Gamma(2m+1)} \prod_{j=0}^{m-1} \frac{\Gamma(2(m+j))}{\Gamma(2(m+j)-2)} \\
&= \frac{2\Gamma(4m-4)}{\Gamma(2m-1)^2 \Gamma(2m+1)} \frac{\Gamma(4m-2)}{\Gamma(2m-2)}.
\end{aligned}$$

To bound $\frac{2\Gamma(4m-4)\Gamma(4m-2)}{\Gamma(2m-1)^2 \Gamma(2m)\Gamma(2m-2)}$ from above we use the binomial relation (see e.g. [3, Eq. 6.1.21]) for the Gamma function

$$\binom{z}{w} = \frac{\Gamma(z+1)}{\Gamma(w+1)\Gamma(z-w+1)} \quad (9.2)$$

and the inequality (see e.g. [13, Eq. C.5])

$$\left(\frac{z}{w}\right)^w \leq \binom{z}{w} \leq \left(\frac{ez}{w}\right)^w \quad (9.3)$$

to obtain

$$\begin{aligned}
\frac{\Gamma(4m-4)\Gamma(4m-2)}{\Gamma(2m-1)^2 \Gamma(2m)\Gamma(2m-2)} &= \frac{\Gamma(4m-4)}{\Gamma(2m-1)\Gamma(2m-2)} \cdot \frac{\Gamma(4m-2)}{\Gamma(2m-1)\Gamma(2m)} \\
&= \binom{4m-5}{2m-2} \binom{4m-3}{2m-2} \\
&= \binom{4m-5}{2m-2} \binom{4m-3}{2m-1} \\
&\leq \left(e \frac{4m-5}{2m-2}\right)^{2m-2} \left(e \frac{4m-3}{2m-1}\right)^{2m-1} \\
&\leq (2e)^{2m-2} (2e)^{2m-1} \\
&= (2e)^{4m-3} \\
&\leq \frac{1}{2} (2e)^{4m-2}.
\end{aligned}$$

Recalling (9.1) establishes the assertion for $n = 2m - 1$.

Now consider the even case $n = 2m$. We utilize the bijective principal representation $\phi_{el}^p : \Lambda^m \times T^m \rightarrow \text{Int}(M^{2m})$ defined in (5.7) and, proceeding as in the odd case, we obtain

$$\frac{\text{Vol}(M_\delta^{2m})}{\text{Vol}(M^{2m})} \geq 1 - 2m\delta \frac{S_{m-1}(3, 1, 2)}{S_m(3, 1, 2)}$$

Using Selberg's formulas (6.1) we compute

$$\begin{aligned}
\frac{S_{m-1}(3, 1, 2)}{S_m(3, 1, 2)} &= \frac{\prod_{j=0}^{m-2} \frac{\Gamma(1+2j)\Gamma(3+2j)^2}{2\Gamma(2(m+j))}}{\prod_{j=0}^{m-1} \frac{\Gamma(1+2j)\Gamma(3+2j)^2}{2\Gamma(2(m+j)+2)}} \\
&= \frac{2\Gamma(4m-2)}{\Gamma(2m-1)^2\Gamma(2m+1)} \frac{\prod_{j=0}^{m-1} \frac{\Gamma(1+2j)\Gamma(3+2j)^2}{2\Gamma(2(m+j))}}{\prod_{j=0}^{m-1} \frac{\Gamma(1+2j)\Gamma(3+2j)^2}{2\Gamma(2(m+j)+2)}} \\
&= \frac{2\Gamma(4m-2)}{\Gamma(2m-1)^2\Gamma(2m+1)} \prod_{j=0}^{m-1} \frac{\Gamma(2(m+j)+2)}{\Gamma(2(m+j))} \\
&= \frac{2\Gamma(4m-2)}{\Gamma(2m-1)^2\Gamma(2m+1)} \frac{\Gamma(4m)}{\Gamma(2m)}.
\end{aligned}$$

We will now use the Beta function

$$B(a, b) := \int_0^1 t^{a-1}(1-t)^{b-1}dt, \quad a > 0, b > 0, \quad (9.4)$$

and the identity

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad (9.5)$$

see e.g. [3, Pg. 258], where Γ is the gamma function.

To bound $\frac{2\Gamma(4m-2)\Gamma(4m)}{\Gamma(2m-1)^2\Gamma(2m)^2}$ from above we use the inequality

$$B(a, a) = \frac{\Gamma(a)^2}{\Gamma(2a)} \geq \frac{4}{a} 2^{-2a}$$

from Proposition 10.2 to obtain

$$\begin{aligned}
\frac{\Gamma(4m-2)\Gamma(4m)}{\Gamma(2m-1)^2\Gamma(2m)^2} &= \frac{1}{B(2m-1, 2m-1)} \frac{1}{B(2m, 2m)} \\
&\leq \frac{2m-1}{4} 2^{4m-2} \frac{2m}{4} 2^{4m} \\
&\leq \frac{2m}{4} 2^{4m-2} \frac{2m}{4} 2^{4m} \\
&\leq \frac{1}{16} m^2 2^{8m}
\end{aligned}$$

Finally, we apply the inequality

$$m^2 \leq 8\left(\frac{e}{2}\right)^{4m}$$

from Proposition 10.1 to conclude that

$$\begin{aligned}
\frac{2\Gamma(4m-2)\Gamma(4m)}{\Gamma(2m-1)^2\Gamma(2m)^2} &\leq \frac{1}{8} m^2 2^{8m} \\
&\leq \left(\frac{e}{2}\right)^{4m} 2^{8m} \\
&= (2e)^{4m}
\end{aligned}$$

thus establishing the assertion for $n = 2m$.

9.5 Proof of Lemma 4.4

According to Chang, Kemperman, Studden [11, Thm. 1.3] one can show, using Skibinsky's canonical coordinates for the moment problem [39], that the uniform distribution on M^n marginalizes to a Beta distribution corresponding to $B(n, n)$ (see (9.4) and (9.5)) on the first moment. Consequently,

$$\begin{aligned} \frac{\text{Vol}(q \in M^n : q_1 \in [1 - \delta, 1])}{\text{Vol}(M^n)} &= \frac{1}{B(n, n)} \int_{1-\delta}^1 t^{n-1} (1-t)^{n-1} dt \\ &= I_\delta(n, n) \end{aligned}$$

where $I_\delta(n, n)$ is the Incomplete Beta function (see e.g. [3, Pg. 258]). Using the binomial relations (9.2) and (9.3), for the upper bound we obtain

$$\begin{aligned} I_\delta(n, n) &:= \frac{1}{B(n, n)} \int_0^\delta t^{n-1} (1-t)^{n-1} dt \\ &\leq \frac{1}{B(n, n)} \int_0^\delta t^{n-1} dt \\ &= \frac{\delta^n}{nB(n, n)} \\ &= \delta^n \frac{\Gamma(2n)}{n\Gamma(n)^2} \\ &= \delta^n \binom{2n-1}{n} \\ &\leq \delta^n \left(e \frac{2n-1}{n} \right)^n \\ &\leq \delta^n (2e)^n \end{aligned}$$

and for the lower bound

$$\begin{aligned} I_\delta(n, n) &:= \frac{1}{B(n, n)} \int_0^\delta t^{n-1} (1-t)^{n-1} dt \\ &\geq \frac{1}{B(n, n)} (1-\delta)^{n-1} \int_0^\delta t^{n-1} dt \\ &= \frac{1}{nB(n, n)} (1-\delta)^{n-1} \delta^n \\ &= \binom{2n-1}{n-1} (1-\delta)^{n-1} \delta^n \\ &\geq \left(\frac{2n-1}{n-1} \right)^{n-1} (1-\delta)^{n-1} \delta^n \\ &\geq 2^{n-1} (1-\delta)^{n-1} \delta^n \\ &\geq \delta^n \end{aligned}$$

where the assumption $\delta \leq \frac{1}{2}$ was used in the last step.

9.6 Proof of Proposition 5.5

The following identity of Karlin and Shapley [25, Proof of Thm. 6.2] will be useful in all the Jacobian determinant calculations of this paper: For $t_1 < s_1 < \dots < t_m < s_m$, we have

$$\frac{\partial^m}{\partial s_1 \dots \partial s_m} \Delta(t_1, s_1, \dots, t_m, s_m) |_{(s_1, \dots, s_m) = (t_1, \dots, t_m)} = \Delta_m^4(t). \quad (9.6)$$

We can develop the upper and lower configurations simultaneously, by introducing a point $t_0 \in \{0, 1\}$ and representations ϕ_{t_0} where when $t_0 = 0$ we have $\phi_0 = \phi_{ol}^c$ defined in (5.15) and when $t_0 = 1$ we have $\phi_1 = \phi_{ou}^c$ defined in (5.16). So, let us use this notation and a change of indices, and consider the two maps

$$\phi_{t_0} : \Lambda^m \times T^{m-1} \rightarrow \text{Int}(M^{2m-1}), \quad t_0 = 0, 1$$

defined by

$$\begin{aligned} \phi_{t_0}(\lambda, t) &= \Psi \left(\sum_{j=0}^{m-1} \lambda_j \delta_{t_j} + (1 - \sum_{j=0}^{m-1} \lambda_j) \delta_{t_*} \right) \\ &= \left(\sum_{j=0}^{m-1} \lambda_j t_j^i + (1 - \sum_{j=0}^{m-1} \lambda_j) t_*^i \right)_{i=1}^{2m-1}. \end{aligned} \quad (9.7)$$

In this notation, Proposition 5.5 becomes

Proposition 9.1. *For $t_0 = 0, 1$ we have*

$$|\det(d\phi_{t_0})|(\lambda, t) = \mathcal{J}_{t_0}(t) \prod_{j=1}^{m-1} \lambda_j$$

where

$$\mathcal{J}_{t_0}(t) = |t_0 - t_*| \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} (t_j - t_0)^2 \cdot \Delta_{m-1}^4(t)$$

The differential of ϕ_{t_0} is determined by

$$\frac{\partial \phi_{t_0}^i}{\partial \lambda_j} = t_j^i - t_*^i, \quad j = 0, \dots, m-1$$

and

$$\frac{\partial \phi_{t_0}^i}{\partial t_j} = i \lambda_j t_j^{i-1} \quad j = 1, \dots, m-1$$

for $i = 1, \dots, 2m-1$, from which we conclude that

$$|\det(d\phi_{t_0})| = |\mathcal{J}_{t_0}| \prod_{j=1}^{m-1} \lambda_j$$

where

$$\begin{aligned} \mathcal{J}_{t_0} &= \begin{vmatrix} t_0 - t_* & t_1 - t_* & 1 & \cdots & t_{m-1} - t_* & 1 \\ t_0^2 - t_*^2 & t_1^2 - t_*^2 & 2t_1 & \cdots & t_{m-1}^2 - t_*^2 & 2t_{m-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ t_0^{2m-1} - t_*^{2m-1} & t_1^{2m-1} - t_*^{2m-1} & (2m-1)t_1^{2m-2} & \cdots & t_{m-1}^{2m-1} - t_*^{2m-1} & (2m-1)t_{m-1}^{2m-2} \end{vmatrix} \\ &= \begin{vmatrix} 1 & 1 & 1 & 0 & \cdots & 1 & 0 \\ t_* & t_0 & t_1 & 1 & \cdots & t_{m-1} & 1 \\ t_*^2 & t_0^2 & t_1^2 & 2t_1 & \cdots & t_{m-1}^2 & 2t_{m-1} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ t_*^{2m-1} & t_0^{2m-1} & t_1^{2m-1} & (2m-1)t_1^{2m-2} & \cdots & t_{m-1}^{2m-1} & (2m-1)t_{m-1}^{2m-2} \end{vmatrix} \end{aligned}$$

To evaluate \mathcal{J}_{t_0} for $t_0 = 0, 1$, let s_1, \dots, s_{m-1} satisfy $t_j < s_j < t_{j+1}, j = 1, \dots, m-1$ and define the Vandermonde determinant

$$\mathcal{J}(s_1, \dots, s_{m-1}) := \begin{vmatrix} 1 & 1 & 1 & s_1^0 & \cdots & 1 & s_{m-1}^0 \\ t_* & t_0 & t_1 & s_1 & \cdots & t_{m-1} & s_{m-1} \\ t_*^2 & t_0^2 & t_1^2 & s_1^2 & \cdots & t_{m-1}^2 & s_{m-1}^2 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ t_*^{2m-1} & t_0^{2m-1} & t_1^{2m-1} & s_1^{2m-1} & \cdots & t_{m-1}^{2m-1} & s_{m-1}^{2m-1} \end{vmatrix}$$

and observe that the multilinearity of the determinant shows that

$$\mathcal{J}_{t_0} = \frac{\partial^{m-1}}{\partial s_1 \cdots \partial s_{m-1}} \mathcal{J}(s_1, \dots, s_{m-1})|_{(s_1, \dots, s_{m-1})=(t_1, \dots, t_{m-1})}$$

To evaluate this differentiation, observe that

$$\begin{aligned} \mathcal{J}(s_1, \dots, s_{m-1}) &= \Delta(t_*, t_0, t_1, s_1, \dots, t_{m-1}, s_{m-1}) \\ &= (t_0 - t_*) \left(\prod_{j=1}^{m-1} (t_j - t_*)(s_j - t_*) \right) \left(\prod_{j=1}^{m-1} (t_j - t_0)(s_j - t_0) \right) \\ &\quad \Delta(t_1, s_1, \dots, t_{m-1}, s_{m-1}), \end{aligned}$$

from which we conclude that

$$\begin{aligned} \mathcal{J}_{t_0} &= \frac{\partial^{m-1}}{\partial s_1 \cdots \partial s_{m-1}} \mathcal{J}(s_1, \dots, s_{m-1})|_{(s_1, \dots, s_{m-1})=(t_1, \dots, t_{m-1})} \\ &= (t_0 - t_*) \cdot \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} (t_j - t_0)^2 \cdot \frac{\partial^{m-1}}{\partial s_1 \cdots \partial s_{m-1}} \\ &\quad \Delta(t_1, s_1, \dots, t_{m-1}, s_{m-1})|_{(s_1, \dots, s_{m-1})=(t_1, \dots, t_{m-1})} \end{aligned}$$

Using the identity (9.6) we conclude that

$$\mathcal{J}_{t_0} = (t_0 - t_*) \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} (t_j - t_0)^2 \Delta_{m-1}^4(t)$$

thereby proving Proposition 9.1 and therefore Proposition 5.5.

9.7 Proof of Proposition 5.6

To simplify notation, let $\phi_l := \phi_{el}^c$ defined in (5.18) and $\phi_u := \phi_{eu}^c$ defined in (5.19). We begin with the lower representation ϕ_l . The differential of ϕ_l is determined by

$$\frac{\partial \phi_l^i}{\partial \lambda_j} = t_j^i - t_*^i, \quad j = 1, \dots, m,$$

and

$$\frac{\partial \phi_l^i}{\partial t_j} = i \lambda_j t_j^{i-1} \quad j = 1, \dots, m$$

for $i = 1, \dots, 2m$, from which we conclude that

$$|\det(d\phi_l)| = |\mathcal{J}_l| \prod_{j=1}^m \lambda_j$$

where

$$\begin{aligned} \mathcal{J}_l &= \begin{vmatrix} t_1 - t_* & 1 & \cdots & t_m - t_* & 1 \\ t_1^2 - t_*^2 & 2t_1 & \cdots & t_m^2 - t_*^2 & 2t_m \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ t_1^{2m} - t_*^{2m} & 2mt_1^{2m-1} & \cdots & t_m^{2m} - t_*^{2m} & 2mt_m^{2m-1} \end{vmatrix} \\ &= \begin{vmatrix} 1 & 1 & 0 & \cdots & 1 & 0 \\ t_* & t_1 & 1 & \cdots & t_m & 1 \\ t_*^2 & t_1^2 & 2t_1 & \cdots & t_m^2 & 2t_m \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ t_*^{2m} & t_1^{2m} & 2mt_1^{2m-1} & \cdots & t_m^{2m} & 2mt_m^{2m-1} \end{vmatrix} \end{aligned}$$

To evaluate \mathcal{J}_l , let s_1, \dots, s_m satisfy $t_j < s_j < t_{j+1}, j = 1, \dots, m$ and define

$$\mathcal{J}(s_1, \dots, s_m) := \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ t_* & t_1 & s_1 & \cdots & t_m & s_m \\ t_*^2 & t_1^2 & s_1^2 & \cdots & t_m^2 & s_m^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ t_*^{2m} & t_1^{2m} & s_1^{2m} & \cdots & t_m^{2m} & s_m^{2m} \end{vmatrix}$$

and observe that the multilinearity of the determinant shows that

$$\mathcal{J}_l = \frac{\partial^m}{\partial s_1 \cdots \partial s_m} \mathcal{J}(s_1, \dots, s_m) \big|_{(s_1, \dots, s_m) = (t_1, \dots, t_m)}$$

To evaluate this differentiation, observe that

$$\mathcal{J}(s_1, \dots, s_m) = \Delta(t_*, t_1, s_1, \dots, t_m, s_m)$$

and, using the recursion relation of the Vandermonde determinant, we obtain

$$\begin{aligned}\mathcal{J}(s_1, \dots, s_m) &= \Delta(t_*, t_1, s_1, \dots, t_m, s_m) \\ &= \prod_{j=1}^m (t_j - t_*)(s_j - t_*) \cdot \Delta(t_1, s_1, \dots, t_m, s_m)\end{aligned}$$

from which we conclude that

$$\begin{aligned}& \frac{\partial^m}{\partial s_1 \dots \partial s_m} \mathcal{J}(s_1, \dots, s_m) \big|_{(s_1, \dots, s_m) = (t_1, \dots, t_m)} \\ &= \prod_{j=1}^m (t_j - t_*)^2 \cdot \frac{\partial^m}{\partial s_1 \dots \partial s_m} \Delta(t_1, s_1, \dots, t_m, s_m) \big|_{(s_1, \dots, s_m) = (t_1, \dots, t_m)} \\ &= \prod_{j=1}^m (t_j - t_*)^2 \cdot \Delta_m^4(t)\end{aligned}$$

and therefore

$$\mathcal{J}_l = \prod_{j=1}^m (t_j - t_*)^2 \cdot \Delta_m^4(t)$$

thus establishing the lower identity.

Now, for the upper representation $\phi_u := \phi_{eu}^c$, the differential of ϕ_u is determined by

$$\frac{\partial \phi_u^i}{\partial \lambda_0} = -t_*^i, ,$$

$$\frac{\partial \phi_u^i}{\partial \lambda_j} = t_j^i - t_*^i, \quad j = 1, \dots, m-1, ,$$

$$\frac{\partial \phi_u^i}{\partial \lambda_m} = 1 - t_*^i, ,$$

and

$$\frac{\partial \phi_u^i}{\partial t_j} = i \lambda_j t_j^{i-1} \quad j = 1, \dots, m-1$$

for $i = 1, \dots, 2m$, from which we conclude that

$$|\det(d\phi_u)| = |\mathcal{J}_u| \prod_{j=1}^{m-1} \lambda_j$$

where

$$\begin{aligned}\mathcal{J}_u &= \begin{vmatrix} -t_* & t_1 - t_* & 1 & \cdots & t_{m-1} - t_* & 1 & 1 - t_* \\ -t_*^2 & t_1^2 - t_*^2 & 2t_1 & \cdots & t_{m-1}^2 - t_*^2 & 2t_{m-1} & 1 - t_*^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ -t_*^{2m} & t_1^{2m} - t_*^{2m} & 2mt_1^{2m-1} & \cdots & t_{m-1}^{2m} - t_*^{2m} & 2mt_{m-1}^{2m-1} & 1 - t_*^{2m} \end{vmatrix} \\ &= \begin{vmatrix} t_* & t_1 & 1 & \cdots & t_{m-1} & 1 & 1 \\ t_*^2 & t_1^2 & 2t_1 & \cdots & t_{m-1}^2 & 2t_{m-1} & 1 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ t_*^{2m} & t_1^{2m} & 2mt_1^{2m-1} & \cdots & t_{m-1}^{2m} & 2mt_{m-1}^{2m-1} & 1 \end{vmatrix}\end{aligned}$$

To evaluate \mathcal{J}_u , let s_1, \dots, s_{m-1} satisfy $t_j < s_j < t_{j+1}, j = 1, \dots, m-1$ and define

$$\mathcal{J}(s_1, \dots, s_{m-1}) := \begin{vmatrix} t_* & t_1 & s_1 & \cdots & t_{m-1} & s_{m-1} & 1 \\ t_*^2 & t_1^2 & s_1^2 & \cdots & t_{m-1}^2 & s_{m-1}^2 & 1 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ t_*^{2m} & t_1^{2m} & s_1^{2m} & \cdots & t_{m-1}^{2m} & s_{m-1}^{2m} & 1 \end{vmatrix}$$

and observe that the multilinearity of the determinant shows that

$$\mathcal{J}_u = \frac{\partial^{m-1}}{\partial s_1 \cdots \partial s_{m-1}} \mathcal{J}(s_1, \dots, s_{m-1})|_{(s_1, \dots, s_{m-1}) = (t_1, \dots, t_{m-1})}$$

To evaluate this differentiation, observe that

$$\mathcal{J}(s_1, \dots, s_{m-1}) = t_* \prod_{j=1}^{m-1} t_j \prod_{j=1}^{m-1} s_j \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & \cdots & 1 & 1 & 1 \\ t_* & t_1 & s_1 & t_2 & s_2 & \cdots & t_{m-1} & s_{m-1} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ t_*^{2m-1} & t_1^{2m-1} & s_1^{2m-1} & t_2^{2m-1} & s_2^{2m-1} & \cdots & t_{m-1}^{2m-1} & s_{m-1}^{2m-1} & 1 \end{vmatrix}$$

That is, we have

$$\mathcal{J}(s_1, \dots, s_{m-1}) = t_* \prod_{j=1}^{m-1} t_j \prod_{j=1}^{m-1} s_j \cdot \Delta(t_*, t_1, s_1, \dots, t_{m-1}, s_{m-1}, 1).$$

We use the recursion relations

$$\Delta(t_*, t_1, s_1, \dots, t_{m-1}, s_{m-1}, 1) = (1 - t_*) \prod_{j=1}^{m-1} (1 - t_j)(1 - s_j) \cdot \Delta(t_*, t_1, s_1, \dots, t_{m-1}, s_{m-1})$$

and

$$\Delta(t_*, t_1, s_1, \dots, t_{m-1}, s_{m-1}) = \prod_{j=1}^{m-1} (t_j - t_*)(s_j - t_j^*) \cdot \Delta(t_1, s_1, \dots, t_{m-1}, s_{m-1})$$

to obtain

$$\mathcal{J}(s_1, \dots, s_{m-1}) = t_*(1-t_*) \prod_{j=1}^{m-1} t_j(1-t_j)s_j(1-s_j) \cdot \prod_{j=1}^{m-1} (t_j - t_*)(s_j - t_*) \cdot \Delta(t_1, s_1, \dots, t_{m-1}, s_{m-1}).$$

Consequently, the identity (9.6) implies

$$\begin{aligned} \mathcal{J}_u &= \frac{\partial^{m-1}}{\partial s_1 \cdots \partial s_{m-1}} \mathcal{J}(s_1, \dots, s_{m-1}) \Big|_{(s_1, \dots, s_{m-1})=(t_1, \dots, t_{m-1})} \\ &= t_*(1-t_*) \prod_{j=1}^{m-1} t_j^2(1-t_j)^2 \prod_{j=1}^{m-1} (t_j - t_*)^2 \cdot \frac{\partial^{m-1}}{\partial s_1 \cdots \partial s_{m-1}} \\ &\quad \Delta(t_*, t_1, s_1, \dots, t_{m-1}, s_{m-1}) \Big|_{(s_1, \dots, s_{m-1})=(t_1, \dots, t_{m-1})} \\ &= t_*(1-t_*) \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} t_j^2(1-t_j)^2 \cdot \Delta_{m-1}^4(t) \end{aligned}$$

establishing the upper identity and thus completing the proof.

9.8 Proof of Proposition 6.1

For the first assertion, let $n = 2m - 1$. As in the proof of Proposition 5.5 we find it convenient to analyze the upper and lower configurations simultaneously, by introducing a point $t_0 \in \{0, 1\}$ and the volume filling representations $\phi_{t_0}, t_0 \in \{0, 1\}$ where when $t_0 = 0$ we have $\phi_0 = \phi_{ol}^c$ defined in (5.15) and when $t_0 = 1$ we have $\phi_1 = \phi_{ou}^c$ defined in (5.16). In this notation, from (9.7) we have

$$\begin{aligned} \phi_{t_0}^i(\lambda_0, \dots, \lambda_{m-1}; t_1, \dots, t_{m-1}) &= \sum_{j=0}^{m-1} \lambda_j t_j^i + (1 - \sum_{j=0}^{m-1} \lambda_j) t_*^i \\ &= \sum_{j=0}^{m-1} \lambda_j (t_j^i - t_*^i) + t_*^i \\ &= \lambda_0 (t_0^i - t_*^i) + \sum_{j=1}^{m-1} \lambda_j (t_j^i - t_*^i) + t_*^i \end{aligned} \quad (9.8)$$

for $t_0 = 0, 1$ and Proposition 9.1 expresses the Jacobian determinants as

$$|\det(d\phi_{t_0})|(\lambda, t) = \mathcal{J}_{t_0}(t) \prod_{j=1}^{m-1} \lambda_j$$

where

$$\mathcal{J}_{t_0}(t) = |t_0 - t_*| \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} (t_j - t_0)^2 \cdot \Delta_{m-1}^4(t). \quad (9.9)$$

In this notation, the modified change of variables formula (5.28) becomes

$$\int_{M^{2m-1}} q_i = \sum_{t_0=1,2} \int_{\Lambda^m \times T^{m-1}} \phi_{t_0}^i |d\phi_{t_0}| \quad (9.10)$$

for $i \geq 0$. Therefore, we conclude that

$$\begin{aligned} \int_{M^{2m-1}} q_i &= \sum_{t_0=1,2} \int_{\Lambda^m \times T^{m-1}} \phi_{t_0}^i |d\phi_{t_0}| \\ &= \sum_{t_0=1,2} \int_{\Lambda^m \times T^{m-1}} \phi_{t_0}^i \left(\prod_{j=1}^{m-1} \lambda_j \right) \mathcal{J}_{t_0} \\ &= \sum_{t_0=1,2} \int_{T^{m-1}} \left(\int_{\Lambda^m} \phi_{t_0}^i \prod_{j=1}^{m-1} \lambda_j d\lambda \right) \mathcal{J}_{t_0} \end{aligned} \quad (9.11)$$

Performing the Λ^m integration, using the identities $\int_{\Lambda^m} \lambda_{m-1}^2 \prod_{i=1}^{m-2} \lambda_i d\lambda = \frac{2}{(2m)!}$, $\int_{\Lambda^m} \prod_{i=1}^m \lambda_i d\lambda = \frac{1}{(2m)!}$, and $\int_{\Lambda^m} \prod_{i=1}^{m-1} \lambda_i d\lambda = \frac{1}{(2m-1)!}$, we obtain

$$\begin{aligned} \int_{\Lambda^m} \phi_{t_0}^i \prod_{j=1}^{m-1} \lambda_j &= \int_{\Lambda^m} \left(\lambda_0(t_0^i - t_*^i) + \sum_{j=1}^{m-1} \lambda_j(t_j^i - t_*^i) + t_*^i \right) \prod_{j=1}^{m-1} \lambda_j d\lambda \\ &= \frac{1}{(2m)!} (t_0^i - t_*^i) + \frac{2}{(2m)!} \sum_{j=1}^{m-1} (t_j^i - t_*^i) + \frac{1}{(2m-1)!} t_*^i \\ &= \frac{1}{(2m)!} t_0^i + \frac{2}{(2m)!} \sum_{j=1}^{m-1} t_j^i + \frac{1}{(2m)!} t_*^i. \end{aligned}$$

Consequently, for $i \geq 1$, we have

$$\begin{aligned} \int_{M^{2m-1}} q_i &= \sum_{t_0=1,2} \int_{T^{m-1}} \left(\int_{\Lambda^m} \phi_{t_0}^i \prod_{j=1}^{m-1} \lambda_j d\lambda \right) \mathcal{J}_{t_0} \\ &= \sum_{t_0=1,2} \int_{T^{m-1}} \left(\frac{1}{(2m)!} t_0^i + \frac{2}{(2m)!} \sum_{j=1}^{m-1} t_j^i + \frac{1}{(2m)!} t_*^i \right) \mathcal{J}_{t_0} \\ &= \frac{1}{(2m)!} \int_{T^{m-1}} \mathcal{J}_1 + \frac{1}{(2m)!} t_*^i \int_{T^{m-1}} (\mathcal{J}_0 + \mathcal{J}_1) + \frac{2}{(2m)!} \int_{T^{m-1}} \left(\sum_{j=1}^{m-1} t_j^i \right) (\mathcal{J}_0 + \mathcal{J}_1) \end{aligned}$$

and, for $i = 0$

$$\int_{M^{2m-1}} q_0 = \frac{1}{(2m-1)!} \int_{T^{m-1}} (\mathcal{J}_0 + \mathcal{J}_1) \quad (9.12)$$

which we already knew from (6.7). Combining the two, we obtain for $i \geq 0$

$$\begin{aligned} \int_{M^{2m-1}} q_i &= \frac{\delta_0(i)}{(2m)!} \int_{T^{m-1}} \mathcal{J}_0 + \frac{1}{(2m)!} \int_{T^{m-1}} \mathcal{J}_1 + \frac{1}{(2m)!} t_*^i \int_{T^{m-1}} (\mathcal{J}_0 + \mathcal{J}_1) \\ &+ \frac{2}{(2m)!} \int_{T^{m-1}} \left(\sum_{j=1}^{m-1} t_j^i \right) (\mathcal{J}_0 + \mathcal{J}_1) \end{aligned} \quad (9.13)$$

and the substitution of the volume equality (9.12) (that is, (6.7)) yields the assertion in the odd case.

For the even case, let $n = 2m$, and let us simplify notation by denoting the volume filling representations by $\phi_1 := \phi_{el}^c$ and $\phi_2 := \phi_{eu}^c$ defined in (5.18) and (5.19) so that, in this notation,

$$\phi_1 : \Lambda^m \times T^m \rightarrow \text{Int}(M^{2m})$$

is defined by

$$\phi_1(\lambda_1, \dots, \lambda_m; t_1, \dots, t_m) = \left(\sum_{j=1}^m \lambda_j t_j^i + (1 - \sum_{j=1}^m \lambda_j) t_*^i \right)_{i=1}^{2m}$$

and

$$\phi_2 : \Lambda^{m+1} \times T^{m-1} \rightarrow \text{Int}(M^{2m})$$

by

$$\phi_2(\lambda_0, \dots, \lambda_m; t_1, \dots, t_{m-1}) = \left(\sum_{j=1}^{m-1} \lambda_j t_j^i + \lambda_m + (1 - \sum_{j=0}^m \lambda_j) t_*^i \right)_{i=1}^{2m}.$$

From Proposition 5.6 we have

$$\begin{aligned} |\det(d\phi_1)(\lambda, t)| &= \mathcal{J}_1(t) \prod_{j=1}^m \lambda_j \\ |\det(d\phi_2)(\lambda, t)| &= \mathcal{J}_2(t) \prod_{j=1}^{m-1} \lambda_j \end{aligned}$$

where

$$\begin{aligned} \mathcal{J}_1(t) &= \prod_{j=1}^m (t_j - t_*)^2 \cdot \Delta_m^4(t) \\ \mathcal{J}_2(t) &= t_* (1 - t_*) \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} t_j^2 (1 - t_j)^2 \cdot \Delta_{m-1}^4(t). \end{aligned}$$

In this notation, the modified change of variable formula (5.28) becomes

$$\int_{M^{2m}} q_i = \int_{\Lambda^m \times T^m} \phi_1^i |d\phi_1| + \int_{\Lambda^{m+1} \times T^{m-1}} \phi_2^i |d\phi_2|. \quad (9.14)$$

We evaluate the two integrals in (9.14) by

$$\begin{aligned}\int_{\Lambda^m \times T^m} \phi_1^i |d\phi_1| &= \int_{\Lambda^m \times T^m} \phi_1^i \left(\prod_{j=1}^m \lambda_j \right) \mathcal{J}_1 \\ &= \int_{T^m} \left(\int_{\Lambda^m} \phi_1^i \prod_{j=1}^m \lambda_j d\lambda \right) \mathcal{J}_1\end{aligned}$$

and

$$\begin{aligned}\int_{\Lambda^{m+1} \times T^{m-1}} \phi_2^i |d\phi_1| &= \int_{\Lambda^{m+1} \times T^{m-1}} \phi_2^i \left(\prod_{j=1}^{m-1} \lambda_j \right) \mathcal{J}_2 \\ &= \int_{T^{m-1}} \left(\int_{\Lambda^{m+1}} \phi_1^i \prod_{j=1}^{m-1} \lambda_j d\lambda \right) \mathcal{J}_2\end{aligned}$$

Performing the Λ^{m+1} and Λ^m integrations, using the identities $\int_{\Lambda^m} \lambda_1 \prod_{j=1}^m \lambda_j d\lambda = \frac{2}{(2m+1)!}$, $\int_{\Lambda^m} \lambda_{m-2}^2 \prod_{i=1}^{m-3} \lambda_i d\lambda = \frac{2}{(2m-1)!}$, $\int_{\Lambda^m} \prod_{j=1}^{m-2} \lambda_j d\lambda = \frac{1}{(2m-2)!}$, $\int_{\Lambda^m} \lambda_{m-1}^2 \prod_{i=1}^{m-2} \lambda_i d\lambda = \frac{2}{(2m)!}$, $\int_{\Lambda^m} \prod_{i=1}^m \lambda_i d\lambda = \frac{1}{(2m)!}$, and $\int_{\Lambda^m} \prod_{i=1}^{m-1} \lambda_i d\lambda = \frac{1}{(2m-1)!}$, we obtain for $i \geq 1$

$$\begin{aligned}\int_{\Lambda^m} \phi_1^i \prod_{j=1}^m \lambda_j d\lambda &= \int_{\Lambda^m} \left(\sum_{j=1}^m \lambda_j t_j^i + (1 - \sum_{j=1}^m \lambda_j) t_*^i \right) \prod_{j=1}^m \lambda_j d\lambda \\ &= \int_{\Lambda^m} \left(\sum_{j=1}^m \lambda_j (t_j^i - t_*^i) + t_*^i \right) \prod_{j=1}^m \lambda_j d\lambda \\ &= \frac{2}{(2m+1)!} \sum_{j=1}^m (t_j^i - t_*^i) + \frac{1}{(2m)!} t_*^i \\ &= \frac{2}{(2m+1)!} \sum_{j=1}^m t_j^i + \frac{1}{(2m+1)!} t_*^i.\end{aligned}$$

and

$$\begin{aligned}\int_{\Lambda^{m+1}} \phi_2^i \prod_{j=1}^{m-1} \lambda_j d\lambda &= \int_{\Lambda^{m+1}} \left(\sum_{j=1}^{m-1} \lambda_j t_j^i + \lambda_m + (1 - \sum_{j=0}^m \lambda_j) t_*^i \right) \prod_{j=1}^{m-1} \lambda_j d\lambda \\ &= \int_{\Lambda^{m+1}} \left(\sum_{j=1}^{m-1} \lambda_j (t_j^i - t_*^i) + \lambda_m (1 - t_*^i) + (1 - \lambda_0) t_*^i \right) \prod_{j=1}^{m-1} \lambda_j d\lambda \\ &= \frac{2}{(2m+1)!} \sum_{j=1}^{m-1} (t_j^i - t_*^i) + \frac{1}{(2m+1)!} (1 - t_*^i) + \left(\frac{1}{(2m)!} - \frac{1}{(2m+1)!} \right) t_*^i \\ &= \frac{2}{(2m+1)!} \sum_{j=1}^{m-1} t_j^i + \frac{1}{(2m+1)!} + \frac{1}{(2m+1)!} t_*^i.\end{aligned}$$

For $i = 0$, (9.14) implies

$$Vol(M^{2m}) = \frac{1}{(2m)!} \int_{T^m} \mathcal{J}_1 + \frac{1}{(2m)!} \int_{T^{m-1}} \mathcal{J}_2$$

so that for $i \geq 1$ we have

$$\begin{aligned} & \int_{M^{2m}} q_i \\ &= \int_{\Lambda^m \times T^m} \phi_1^i |d\phi_1| + \int_{\Lambda^{m+1} \times T^{m-1}} \phi_2^i |d\phi_2| \\ &= \int_{T^m} \left(\frac{2}{(2m+1)!} \sum_{j=1}^m t_j^i + \frac{1}{(2m+1)!} t_*^i \right) \mathcal{J}_1 \\ &+ \int_{T^{m-1}} \left(\frac{2}{(2m+1)!} \sum_{j=1}^{m-1} t_j^i + \frac{1}{(2m+1)!} + \frac{1}{(2m+1)!} t_*^i \right) \mathcal{J}_2 \\ &= \int_{T^m} \left(\frac{2}{(2m+1)!} \sum_{j=1}^m t_j^i \right) \mathcal{J}_1 \\ &+ \int_{T^{m-1}} \left(\frac{2}{(2m+1)!} \sum_{j=1}^{m-1} t_j^i + \frac{1}{(2m+1)!} \right) \mathcal{J}_2 \\ &+ \frac{t_*^i}{2m+1} Vol(M^{2m}) \end{aligned}$$

so that for $i \geq 1$ we conclude

$$\begin{aligned} & \int_{M^{2m}} q_i - \frac{t_*^i}{2m+1} Vol(M^{2m}) \\ &= \frac{1}{(2m+1)!} \int_{T^m} \left(2 \sum_{j=1}^m t_j^i \right) \mathcal{J}_1 + \frac{1}{(2m+1)!} \int_{T^{m-1}} \left(2 \sum_{j=1}^{m-1} t_j^i + 1 \right) \mathcal{J}_2. \end{aligned}$$

Combining with the result (9.15) for $i = 0$ we conclude

$$\begin{aligned} & \int_{M^{2m}} q_i - \frac{t_*^i}{2m+1} Vol(M^{2m}) \\ &= \frac{2}{(2m+1)!} \int_{T^m} \sum_{j=1}^m t_j^i \mathcal{J}_1 + \frac{2}{(2m+1)!} \int_{T^{m-1}} \sum_{j=1}^{m-1} t_j^i \mathcal{J}_2 + \frac{1 + \delta_0(i)}{(2m+1)!} \int_{T^{m-1}} \mathcal{J}_2 \end{aligned}$$

establishing the assertion in the even case.

9.9 Proof of Theorem 7.1

Recall the identity $\mathcal{J}_{ol}^c(0, t) \equiv 0$. Then subtracting the volume identity (6.7)

$$Vol(M^{2m-1}) = \frac{1}{(2m-1)!(m-1)!} \int_{I^{m-1}} (\mathcal{J}_{ol}^c(t_*, t) + \mathcal{J}_{ou}^c(t_*, t)) dt \quad (9.15)$$

from itself evaluated at $t_* = 0$, we conclude that

$$\int_{I^{m-1}} (\mathcal{J}_{ol}^c(t_*, t) + \mathcal{J}_{ou}^c(t_*, t) - \mathcal{J}_{ou}^c(0, t)) dt \equiv 0, \quad (9.16)$$

that is,

$$\int_{I^{m-1}} \mathcal{H}(t_*, t) dt \equiv 0.$$

We now do the same subtraction for all the moments. To that end, recall the convention $0^0 = 1$, and observe that, for $i \geq 0$, the identity

$$\begin{aligned} & \int_{M^{2m-1}} q_i - \frac{t_*^i}{2m} \text{Vol}(M^{2m-1}) \\ &= \frac{\delta_0(i)}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ol}^c(t_*, t) dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ou}^c(t_*, t) dt \\ &+ \frac{2}{(2m)!(m-1)!} \int_{I^{m-1}} \Sigma t^i (\mathcal{J}_{ol}^c(t_*, t) + \mathcal{J}_{ou}^c(t_*, t)) dt \end{aligned} \quad (9.17)$$

from Proposition 6.1, evaluated at $t_* = 0$ becomes

$$\begin{aligned} & \int_{M^{2m-1}} q_i - \frac{\delta_0(i)}{2m} \text{Vol}(M^{2m-1}) \\ &= \frac{\delta_0(i)}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ol}^c(0, t) dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ou}^c(0, t) dt \\ &+ \frac{2}{(2m)!(m-1)!} \int_{I^{m-1}} \Sigma t^i (\mathcal{J}_{ol}^c(0, t) + \mathcal{J}_{ou}^c(0, t)) dt. \end{aligned}$$

Subtracting from (9.17), using the identity $\mathcal{J}_{ol}^c(0, t) \equiv 0$, we obtain

$$\begin{aligned} & -\frac{t_*^i}{2m} \text{Vol}(M^{2m-1}) + \frac{\delta_0(i)}{2m} \text{Vol}(M^{2m-1}) \\ &= \frac{\delta_0(i)}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ol}^c(t_*, t) dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} (\mathcal{J}_{ou}^c(t_*, t) - \mathcal{J}_{ou}^c(0, t)) dt \\ &- \frac{2}{(2m)!(m-1)!} \int_{I^{m-1}} \Sigma t^i \mathcal{H}(t_*, t) dt \end{aligned}$$

and applying the volume identity (9.15) we obtain

$$\begin{aligned} & -\frac{t_*^i}{2m} \text{Vol}(M^{2m-1}) \\ &= -\frac{\delta_0(i)}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ou}^c(t_*, t) dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} (\mathcal{J}_{ou}^c(t_*, t) - \mathcal{J}_{ou}^c(0, t)) dt \\ &- \frac{2}{(2m)!(m-1)!} \int_{I^{m-1}} \Sigma t^i \mathcal{H}(t_*, t) dt \end{aligned}$$

and the subtracted volume identity (9.16) we obtain with a change of sign

$$\begin{aligned}
& \frac{t_*^i}{2m} \text{Vol}(M^{2m-1}) \\
&= \frac{\delta_0(i)}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ou}^c(t_*, t) dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ol}^c(t_*, t) dt \\
&+ \frac{2}{(2m)!(m-1)!} \int_{I^{m-1}} \Sigma t^i \mathcal{H}(t_*, t) dt.
\end{aligned} \tag{9.18}$$

Then, if we let $\phi(s) := \sum_{i=0}^{2m-1} \phi_i s^i$ be a polynomial of degree $n = 2m - 1$, summing over each identity in (9.18), we conclude that

$$\begin{aligned}
& \frac{\phi(t_*)}{2m} \text{Vol}(M^{2m-1}) \\
&= \phi_0 \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} (\mathcal{J}_{ol}^c(t_*, t) + \mathcal{J}_{ou}^c(t_*, t)) dt + \sum_{i=1}^{2m-1} \phi_i \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ol}^c(t_*, t) dt \\
&+ \frac{2}{(2m)!(m-1)!} \int_{I^{m-1}} (\Sigma \phi)(t) \mathcal{H}(t_*, t) dt \\
&= \phi_0 \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ou}^c(t_*, t) dt + \sum_{i=0}^{2m-1} \phi_i \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{ol}^c(t_*, t) dt \\
&+ \frac{2}{(2m)!(m-1)!} \int_{I^{m-1}} (\Sigma \phi)(t) \mathcal{H}(t_*, t) dt.
\end{aligned}$$

Since $\phi(0) = \phi_0$ and $\phi(1) = \sum_{i=0}^{2m-1} \phi_i$ the assertion follows by multiplication by $2m$. The even case proceeds in the same way, but since it is a little different we have included it in Section 10.2 in the Appendix.

9.10 Proof of Theorem 8.1

For the first assertion, let $n = 2m - 1$ and consider the integral formula (9.15)

$$\text{Vol}(M^{2m-1}) = \frac{1}{(2m-1)!(m-1)!} \int_{I^{m-1}} (\mathcal{J}_{ol}^c(t_*, t) + \mathcal{J}_{ou}^c(t_*, t)) dt$$

for the volume in terms of the canonical representations. From the definitions

$$\begin{aligned}
\mathcal{J}_{ol}^c(t_*, t) &= t_* \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} t_j^2 \cdot \Delta_{m-1}^4(t) \\
\mathcal{J}_{ou}^c(t_*, t) &= (1 - t_*) \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} (1 - t_j)^2 \cdot \Delta_{m-1}^4(t)
\end{aligned}$$

of Proposition 5.5, we obtain

$$\begin{aligned}\frac{\partial}{\partial t_*} \mathcal{J}_{ol}^c(t_*, t)|_{t_*=0} &= \prod_{j=1}^{m-1} t_j^4 \cdot \Delta_{m-1}^4(t) \\ \frac{\partial}{\partial t_*} \mathcal{J}_{ou}^c(t_*, t)|_{t_*=0} &= -(1 + 2\Sigma t^{-1}) \prod_{j=1}^{m-1} t_j^2 (1 - t_j)^2 \Delta_{m-1}^4(t).\end{aligned}$$

Differentiating the volume formula with respect to t_* at $t_* = 0$, we obtain

$$0 = \frac{1}{(2m-1)!(m-1)!} \int_{I^{m-1}} \left(\frac{\partial}{\partial t_*} \mathcal{J}_{ol}^c(t_*, t)|_{t_*=0} + \frac{\partial}{\partial t_*} \mathcal{J}_{ou}^c(t_*, t)|_{t_*=0} \right) dt$$

and therefore

$$\begin{aligned}\int_{I^{m-1}} \prod_{j=1}^{m-1} t_j^4 \cdot \Delta_{m-1}^4(t) dt &= \int_{I^{m-1}} (1 + 2\Sigma t^{-1}) \prod_{j=1}^{m-1} t_j^2 (1 - t_j)^2 \Delta_{m-1}^4(t) dt \\ &= \int_{I^{m-1}} \prod_{j=1}^{m-1} t_j^2 (1 - t_j)^2 \Delta_{m-1}^4(t) dt \\ &\quad + 2 \int_{I^{m-1}} \Sigma t^{-1} \cdot \prod_{j=1}^{m-1} t_j^2 (1 - t_j)^2 \Delta_{m-1}^4(t) dt\end{aligned}$$

from which we conclude that

$$2 \int_{I^{m-1}} \Sigma t^{-1} \cdot \prod_{j=1}^{m-1} t_j^2 (1 - t_j)^2 \Delta_{m-1}^4(t) dt = S_{m-1}(5, 1, 2) - S_{m-1}(3, 3, 2).$$

Changing $m \mapsto m + 1$ finishes the proof of the first assertion.

For the second assertion, let $n = 2m$ and consider the integral formula (6.8)

$$\text{Vol}(M^{2m}) = \frac{1}{(2m)!m!} \int_{I^m} \mathcal{J}_{el}^c(t_*, t) dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt.$$

From the definitions

$$\begin{aligned}\mathcal{J}_{el}^c(t_*, t) &= \prod_{j=1}^m (t_j - t_*)^2 \cdot \Delta_m^4(t) \\ \mathcal{J}_{eu}^c(t_*, t) &= t_*(1 - t_*) \prod_{j=1}^{m-1} (t_j - t_*)^2 \prod_{j=1}^{m-1} t_j^2 (1 - t_j)^2 \cdot \Delta_{m-1}^4(t)\end{aligned}$$

of Proposition 5.6, we obtain

$$\begin{aligned}\frac{\partial}{\partial t_*} \mathcal{J}_{el}^c(t_*, t)|_{t_*=0} &= -2\Sigma t^{-1} \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t) \\ \frac{\partial}{\partial t_*} \mathcal{J}_{eu}^c(t_*, t)|_{t_*=0} &= \prod_{j=1}^{m-1} t_j^4 (1 - t_j)^2 \cdot \Delta_{m-1}^4(t).\end{aligned}$$

Differentiating the volume formula with respect to t_* at $t_* = 0$, we obtain

$$0 = \frac{1}{(2m)!m!} \int_{I^m} \frac{\partial}{\partial t_*} \mathcal{J}_{el}^c(t_*, t)|_{t_*=0} dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \frac{\partial}{\partial t_*} \mathcal{J}_{eu}^c(t_*, t)|_{t_*=0} dt$$

and therefore we conclude

$$\begin{aligned} 2 \int_{I^m} \Sigma t^{-1} \cdot \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t) dt &= m \int_{I^{m-1}} \prod_{j=1}^{m-1} t_j^4 (1-t_j)^2 \cdot \Delta_{m-1}^4(t) dt \\ &= m S_{m-1}(5, 3, 2), \end{aligned}$$

finishing the proof of the second assertion.

9.11 Proof of Theorem 8.2

First note that the definition (7.1)

$$\mathcal{H}(t_*, t) := \mathcal{J}_{ou}^c(0, t) - \mathcal{J}_{ol}^c(t_*, t) - \mathcal{J}_{ou}^c(t_*, t)$$

and the definitions of \mathcal{J}_{ol}^c and \mathcal{J}_{ou}^c from Proposition 5.5 imply that $\hat{\mathcal{H}}(\cdot, t) \in \Pi^{2m-1}$, $t \in I^{m-1}$. Therefore, it follows from (7.2) that $\hat{\mathcal{H}}(\cdot, t) \in \Pi_0^{2m-1}$, $t \in I^{m-1}$. Now, it follows from Theorem 7.1 that

$$\phi(t_*) = \int_{I^{m-1}} (\Sigma \phi)(t) \hat{\mathcal{H}}(t_*, t) dt, \quad \phi \in \Pi_0^{2m-1}$$

which expanded becomes

$$\phi(t_*) = \sum_{j=1}^{2m-2} p_j(t_*) \int_{I^{m-1}} (\Sigma \phi)(t) h_j(t) dt, \quad \phi \in \Pi_0^{2m-1},$$

in particular, by choosing $\phi := p_k$, $k = 1, \dots, 2m-2$,

$$p_k(t_*) = \sum_{j=1}^{2m-2} p_j(t_*) \int_{I^{m-1}} (\Sigma p_k)(t) h_j(t) dt, \quad k = 1, \dots, 2m-2,$$

from which we conclude

$$\int_{I^{m-1}} \Sigma p_j \cdot h_k = \delta_{jk}, \quad j, k = 1, \dots, 2m-2$$

establishing the assertion. Furthermore, from this and the symmetry of h_k , $k = 1, \dots, 2m-2$ with respect to the action of the symmetric group, we also conclude

$$(m-1) \int_I p_j \cdot \bar{h}_k = \delta_{jk}, \quad j, k = 1, \dots, 2m-2$$

establishing Corollary 8.3.

9.12 Proof of Theorem 8.4

From the orthogonality relation (see e.g. [5, Eq. 12.110])

$$\int_I Q_j Q_k = \frac{1}{2k+1} \frac{(k+2)!}{(k-2)!} \delta_{jk}$$

and the definition

$$\hat{\mathcal{H}} := \frac{1}{\text{Vol}(M^{2m-1})} \frac{2}{(2m-1)!(m-1)!} \mathcal{H}$$

of the scaling of the kernel (7.1)

$$\mathcal{H}(t_*, t) = \mathcal{J}_{ou}^p(t) - \mathcal{J}_{ol}^c(t_*, t) - \mathcal{J}_{ou}^c(t_*, t), \quad (9.19)$$

we can compute the coefficients h_k in the expansion

$$\hat{\mathcal{H}}(t_*, t) = \sum_{j=2}^{2m-1} h_j(t) Q_j(t_*), \quad (t_*, t) \in I \times I^{m-1}$$

as

$$h_k(t) = \frac{1}{\text{Vol}(M^{2m-1})} \frac{2}{(2m-1)!(m-1)!} \frac{(2k+1)(k-2)!}{(k+2)!} \int_I \mathcal{H}(\cdot, t) Q_k(\cdot), \quad k = 2, \dots, 2m-1 \quad (9.20)$$

and then apply Theorem 8.2 to obtain the assertion. To that end, for $k = 2, \dots, 2m-1$, to compute

$$\int_I \mathcal{H}(\cdot, t) Q_k(\cdot)$$

we use the decomposition of (9.19) of \mathcal{H} and compute the values $\int_I \mathcal{J}_{ou}^p(t) Q_k(\cdot)$, $\int_I \mathcal{J}_{ol}^c(\cdot, t) Q_k(\cdot)$, and $\int_I \mathcal{J}_{ou}^c(\cdot, t) Q_k(\cdot)$ separately. For the first term, observe that $\int_I Q_k = 1 + (-1)^k$, $k = 2, \dots, 2m-1$ from Proposition 9.2, so that

$$\int_I \mathcal{J}_{ou}^p(t) Q_k(r) dr = \mathcal{J}_{ou}^p(t) \int_I Q_k(r) dr = (1 + (-1)^k) \mathcal{J}_{ou}^p(t). \quad (9.21)$$

For the second, $\int_I \mathcal{J}_{ol}^c(\cdot, t) Q_k(\cdot)$, we expand $\mathcal{J}_{ol}^c(\cdot, t)$, defined in (5.5), as a polynomial for fixed t and then utilize the values of the integrals $\int_I r^j Q_k(r) dr$ for the monomials r^j , $j = 1, \dots, 2m-1$. To that end, define

$$\tilde{\mathcal{J}}_{ol}^c(\cdot, t) := t * \prod_{j=1}^{m-1} (t_j - t_*)^2 \quad (9.22)$$

so that

$$\mathcal{J}_{ol}^c(\cdot, t) = \tilde{\mathcal{J}}_{ol}^c(\cdot, t) \prod_{j=1}^{m-1} t_j^2 \cdot \Delta_{m-1}^4(t).$$

Then from the definition of \mathcal{J}_{el}^p of (5.4) we have $\mathcal{J}_{el}^p(t) = \prod_{j=1}^{m-1} t_j^2 \cdot \Delta_{m-1}^4(t)$ so that

$$\mathcal{J}_{ol}^c(\cdot, t) = \tilde{\mathcal{J}}_{ol}^c(\cdot, t) \mathcal{J}_{el}^p(t). \quad (9.23)$$

The generating function identity for the elementary symmetric functions e_j is

$$\prod_{i=1}^{m-1} (1 + st_i) = \sum_{j=0}^{m-1} e_j(t) s^j$$

and squaring it we obtain

$$\begin{aligned} \prod_{i=1}^{m-1} (1 + st_i)^2 &= \left(\sum_{j=0}^{m-1} e_j(t) s^j \right)^2 \\ &= \sum_{j=0}^{2m-2} e_j(t, t) s^j. \end{aligned} \quad (9.24)$$

Therefore, by changing $s \mapsto -t_*^{-1}$ we conclude that

$$\prod_{j=1}^{m-1} (t_j - t_*)^2 = \sum_{j=0}^{2m-2} e_{2m-2-j}(t, t) (-1)^j t_*^j \quad (9.25)$$

and therefore

$$t_* \prod_{j=1}^{m-1} (t_j - t_*)^2 = \sum_{j=0}^{2m-2} e_{2m-2-j}(t, t) (-1)^j t_*^{j+1}$$

and, relabeling, we conclude

$$\tilde{\mathcal{J}}_{ol}^c(t_*, t) = - \sum_{j=0}^{2m-1} d_j(t) (-1)^j t_*^j \quad (9.26)$$

where

$$d_j(t) := e_{2m-1-j}(t, t) = \sum_{j_1+j_2=2m-1-j} e_{j_1}(t) e_{j_2}(t), \quad j = 1, \dots, 2m-1 \quad (9.27)$$

and $d_0(t) := 0$. Note that d_j is a symmetric polynomial of degree $2m-1-j$.

In particular, from the definition (9.22) and its resulting polynomial expansion (9.26) we have

$$\tilde{\mathcal{J}}_{ol}^c(1, t) = - \sum_{j=0}^{2m-1} d_j(t) (-1)^j = \prod_{j=1}^{m-1} (1 - t_j)^2. \quad (9.28)$$

The following proposition computes the values of the integrals of the Legendre polynomials against the monomials, and we observe that $\int_I r^j Q_k(r) dr = 1$ for $1 \leq j < k$ and $\int_I r^j Q_k(r) dr = 1$ plus a term when $j \geq k$.

Proposition 9.2. For $k = 2, \dots, 2m - 1$ we have

$$\begin{aligned} \int_I Q_k(r) dr &= 1 + (-1)^k \\ \int_I r^j Q_k(r) dr &= 1, \quad 1 \leq j < k \\ \int_I r^j Q_k(r) dr &= 1 - \frac{(j + k + k^2)\Gamma(j + 2)\Gamma(j)}{\Gamma(j + k + 2)\Gamma(j - k + 1)}, \quad j \geq k. \end{aligned}$$

Therefore, since $d_0 = 0$, we obtain from the polynomial expansion (9.26), Proposition 9.2, and (9.28) that

$$\begin{aligned} \int_I \tilde{\mathcal{J}}_{ol}^c(r, t) Q_k(r) dr &= \sum_{j=k}^{2m-1} (-1)^j a_{jk} d_j(t) - \sum_{j=0}^{2m-1} (-1)^j d_j(t) \\ &= \sum_{j=k}^{2m-1} (-1)^j a_{jk} d_j(t) + \tilde{\mathcal{J}}_{ol}^c(1, t) \end{aligned} \quad (9.29)$$

where we recall the definition

$$a_{jk} := \frac{(j + k + k^2)\Gamma(j + 2)\Gamma(j)}{\Gamma(j + k + 2)\Gamma(j - k + 1)},$$

and note that the lower limit in the summation is k and consequently, the first term is a symmetric polynomial of degree $2m - 1 - k$. Therefore, multiplying by $\mathcal{J}_{el}^p(t) = \prod_{j=1}^{m-1} t_j^2 \cdot \Delta_{m-1}^4(t)$ and using the definition (9.23) and the identity

$$\begin{aligned} \tilde{\mathcal{J}}_{ol}^c(1, t) \mathcal{J}_{el}^p(t) &= \prod_{j=1}^{m-1} (1 - t_j)^2 \prod_{j=1}^{m-1} t_j^2 \cdot \Delta_{m-1}^4(t) \\ &= \mathcal{J}_{ou}^p(t) \end{aligned} \quad (9.30)$$

from the definition $\mathcal{J}_{ou}^p(t) = \prod_{j=1}^{m-1} t_j^2 (1 - t_j)^2 \cdot \Delta_{m-1}^4(t)$ from Proposition (5.3), we conclude that

$$\int_I \mathcal{J}_{ol}^c(r, t) Q_k(r) dr = \mathcal{J}_{el}^p(t) \sum_{j=k}^{2m-1} (-1)^j a_{jk} d_j(t) + \mathcal{J}_{ou}^p(t).$$

Let us designate the negative of the first term

$$\dot{h}_k(t) := \mathcal{J}_{el}^p(t) \sum_{j=k}^{2m-1} (-1)^{j+1} a_{jk} d_j(t) \quad (9.31)$$

so that

$$\int_I \mathcal{J}_{ol}^c(r, t) Q_k(r) dr = -\dot{h}_k(t) + \mathcal{J}_{ou}^p(t). \quad (9.32)$$

For the third term $\int_I \mathcal{J}_{ou}^c(\cdot, t) Q_j(\cdot)$, we utilize the reflection symmetry

$$\mathcal{J}_{ou}^c(t_*, t) = \mathcal{J}_{ol}^c(1 - t_*, 1 - t)$$

of (5.17) and the reflection symmetry (see e.g. [5, Eq. 12.97])

$$Q_k(1 - r) = (-1)^k Q_k(r)$$

of the associate Legendre polynomials to compute the integral $\int_I \mathcal{J}_{ou}^c(\cdot, t) Q_j(\cdot)$ in terms of $\int_I \mathcal{J}_{ol}^c(\cdot, t) Q_j(\cdot)$. That is,

$$\begin{aligned} \int_I \mathcal{J}_{ou}^c(r, t) Q_k(r) dr &= \int_I \mathcal{J}_{ou}^c(1 - r, t) Q_k(1 - r) dr \\ &= \int_I \mathcal{J}_{ol}^c(r, 1 - t) Q_k(1 - r) dr \\ &= (-1)^k \int_I \mathcal{J}_{ol}^c(r, 1 - t) Q_k(r) dr \end{aligned}$$

and therefore

$$\begin{aligned} \int_I \mathcal{J}_{ou}^c(r, t) Q_k(r) dr &= (-1)^k \int_I \mathcal{J}_{ol}^c(r, 1 - t) Q_k(r) dr \\ &= -(-1)^k \dot{h}_k(1 - t) + (-1)^k \mathcal{J}_{ou}^p(1 - t), \end{aligned}$$

so that we conclude

$$\int_I \mathcal{J}_{ou}^c(r, t) Q_k(r) dr = -(-1)^k \dot{h}_k(1 - t) + (-1)^k \mathcal{J}_{ou}^p(1 - t). \quad (9.33)$$

Putting all three terms together using (9.19) and the identities (9.21), (9.32), and (9.33), along with the symmetry

$$\mathcal{J}_{ou}^p(1 - t) = \mathcal{J}_{ou}^p(t),$$

we conclude that

$$\begin{aligned} \int_I \mathcal{H}(r, t) Q_k(r) dr &= \int_I \mathcal{J}_{ou}^p(t) Q_k(r) dr - \int_I \mathcal{J}_{ol}^c(r, t) Q_k(r) dr - \int_I \mathcal{J}_{ou}^c(r, t) Q_k(r) dr \\ &= (1 + (-1)^k) \mathcal{J}_{ou}^p(t) + \dot{h}_k(t) - \mathcal{J}_{ou}^p(t) + (-1)^k \dot{h}_k(1 - t) - (-1)^k \mathcal{J}_{ou}^p(1 - t) \\ &= \dot{h}_k(t) + (-1)^k \dot{h}_k(1 - t). \end{aligned}$$

To finish, consider the functions $\check{h}_k(t) := \dot{h}_k(t) + (-1)^k \dot{h}_k(1 - t)$. It follows from (9.20) that the basis coefficients h_k satisfy

$$h_k = \frac{1}{\text{Vol}(M^{2m-1})} \frac{2}{(2m-1)!(m-1)!} \frac{(2k+1)(k-2)!}{(k+2)!} \check{h}_k, \quad k = 2, \dots, 2m-1.$$

Moreover, Theorem 8.2 implies that $\{\Sigma Q_j, j = 2, \dots, 2m-1\}$ and $\{h_k, k = 2, \dots, 2m-1\}$ are an $L^2(I^{m-1})$ biorthogonal system. It therefore follows that

$$\int_{I^{m-1}} \check{h}_k \Sigma Q_j = \text{Vol}(M^{2m-1})(2m-1)!(m-1)! \frac{(k+2)!}{(4k+2)(k-2)!} \delta_{jk}, \quad j, k = 2, \dots, 2m-1.$$

Moreover, from the symmetry $Q_k(1-r) = (-1)^k Q_k(r)$ and a change of variables we obtain

$$\begin{aligned} \int_{I^{m-1}} \check{h}_k(t) \Sigma Q_j(t) dt &= \int_{I^{m-1}} \dot{h}_k(t) \Sigma Q_j(t) dt + (-1)^k \int_{I^{m-1}} \dot{h}_k(1-t) \Sigma Q_j(t) dt \\ &= \int_{I^{m-1}} \dot{h}_k(t) \Sigma Q_j(t) dt + (-1)^k \int_{I^{m-1}} \dot{h}_k(t) \Sigma Q_j(1-t) dt \\ &= \int_{I^{m-1}} \dot{h}_k(t) \Sigma Q_j(t) dt + (-1)^{j+k} \int_{I^{m-1}} \dot{h}_k(t) \Sigma Q_j(t) dt \\ &= (1 + (-1)^{j+k}) \int_{I^{m-1}} \dot{h}_k(t) \Sigma Q_j(t) dt. \end{aligned}$$

Since $(j+k) \bmod 2 = (j-k) \bmod 2$, the assertion then follows from the definition (9.31) of \dot{h}_k , the identity $\mathcal{J}_{el}^p(t) = \prod_{j=1}^{m-1} t_j^2 \cdot \Delta_{m-1}^4(t)$, and the definition (9.27) of d_k . Moreover, we see that for $(j-k) \bmod 2 = 1$ the vanishing of this integral does not depend on the function \dot{h}_k but is instead a consequence only of the relative parity between Q_j and \dot{h}_k with respect to the operation of reflection.

9.13 Proof of Proposition 9.2

We abuse notation by letting P_k and Q_k denote the Legendre polynomials on the standard set $[-1, 1]$. At the end we will change back to the interval I . We use [20, Eq. 7.127, pg. 771]

$$\int_{-1}^1 (1+x)^j P_k(x) dx = \frac{2^{j+1} \Gamma^2(j+1)}{\Gamma(j+k+2) \Gamma(j-k+1)}, \quad j \geq 0$$

and the definition $Q_k(x) := (1-x^2)P_k''(x)$ and integration by parts. Because of the poles of the Gamma function at 0 and the negative integers, we conclude that

$$\int_{-1}^1 (1+x)^j P_k(x) dx = 0, \quad j < k.$$

Consider the function $\phi(x) := (1+x)^j(1-x^2) = 2(1+x)^{j+1} - (1+x)^{j+2}$, which has the derivatives $\phi'(x) = 2(j+1)(1+x)^j - (j+2)(1+x)^{j+1}$ and $\phi''(x) = 2(j+1)j(1+x)^{j-1} - (j+2)(j+1)(1+x)^j$. Since, for $j \geq 1$, we have $\phi'(-1) = 0$ and $\phi'(1) = -2^{j+1}$,

and $P_k(1) = 1$, we obtain

$$\begin{aligned}
\int_{-1}^1 (1+x)^j Q_k(x) dx &= \int_{-1}^1 (1+x)^j (1-x^2) P_k''(x) dx \\
&= \int_{-1}^1 \phi(x) P_k''(x) dx \\
&= \phi P_k' \Big|_{-1}^1 - \int_{-1}^1 \phi'(x) P_k'(x) dx \\
&= - \int_{-1}^1 \phi'(x) P_k'(x) dx \\
&= -\phi' P_k \Big|_{-1}^1 + \int_{-1}^1 \phi''(x) P_k(x) dx \\
&= 2^{j+1} + \int_{-1}^1 \phi''(x) P_k(x) dx \\
&= 2^{j+1} + 2(j+1)j \int_{-1}^1 (1+x)^{j-1} P_k(x) dx \\
&\quad - (j+2)(j+1) \int_{-1}^1 (1+x)^j P_k(x) dx
\end{aligned}$$

from which we conclude that

$$\int_{-1}^1 (1+x)^j Q_k(x) dx = 2^{j+1}, \quad 1 \leq j < k.$$

For the case $j = 0$, defining $\phi(x) := 1 - x^2$, we instead have $\phi'(-1) = 2$, $\phi'(1) = -2$, $P_k(1) = 1$, and $P_k(-1) = (-1)^k$. Using $\int_{-1}^1 P_k = 0$, $k \geq 1$, we obtain

$$\begin{aligned}
\int_{-1}^1 Q_k(x) dx &= \int_{-1}^1 (1-x^2) P_k''(x) dx \\
&= \int_{-1}^1 \phi(x) P_k''(x) dx \\
&= - \int_{-1}^1 \phi'(x) P_k'(x) dx \\
&= -\phi' P_k \Big|_{-1}^1 + \int_{-1}^1 \phi''(x) P_k(x) dx \\
&= -\phi' P_k \Big|_{-1}^1 \\
&= -(-2P_k(1) - 2P_k(-1)) \\
&= 2(1 + (-1)^k).
\end{aligned}$$

On the other hand, for $j \geq k$ we have

$$\begin{aligned}
& \int_{-1}^1 (1+x)^j Q_k(x) dx - 2^{j+1} \\
&= 2(j+1)j \int_{-1}^1 (1+x)^{j-1} P_k(x) dx - (j+2)(j+1) \int_{-1}^1 (1+x)^j P_k(x) dx \\
&= 2^{j+1}(j+1)j \frac{\Gamma^2(j)}{\Gamma(j+k+1)\Gamma(j-k)} - 2^{j+1}(j+2)(j+1) \frac{\Gamma^2(j+1)}{\Gamma(j+k+2)\Gamma(j-k+1)} \\
&= 2^{j+1}j(j+1)\Gamma^2(j) \left(\frac{1}{\Gamma(j+k+1)\Gamma(j-k)} - j(j+2) \frac{1}{\Gamma(j+k+2)\Gamma(j-k+1)} \right) \\
&= 2^{j+1} \frac{j(j+1)\Gamma^2(j)}{\Gamma(j+k+2)\Gamma(j-k+1)} \left((j+k+1)(j-k) - j(j+2) \right) \\
&= -2^{j+1} \frac{j(j+1)(j+k+k^2)\Gamma^2(j)}{\Gamma(j+k+2)\Gamma(j-k+1)} \\
&= -2^{j+1} \frac{(j+k+k^2)\Gamma(j+2)\Gamma(j)}{\Gamma(j+k+2)\Gamma(j-k+1)}
\end{aligned}$$

and therefore

$$\int_{-1}^1 (1+x)^j Q_k(x) dx = 2^{j+1} \left(1 - \frac{(j+k+k^2)\Gamma(j+2)\Gamma(j)}{\Gamma(j+k+2)\Gamma(j-k+1)} \right).$$

Translating to the unit interval with the map $I \mapsto [-1, 1]$ defined by $x = 2r - 1$ we obtain the assertion.

10 Appendix

10.1 Proof of even case of Lemma 4.3

We utilize the bijective principal representation $\phi_{el}^p : \Lambda^m \times T^m \rightarrow \text{Int}(M^{2m})$ defined in (5.7) and

$$|\det(d\phi_{eu}^p)|(\lambda, t) = \mathcal{J}_{eu}^p(t) \prod_{j=1}^m \lambda_j$$

where

$$\mathcal{J}_{el}^p(t) := \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t)$$

from Proposition 5.4.

Fix $t_* \in (0, 1)$ and let

$$T_\delta^m := \{ (t_1, \dots, t_m) \in T^m : t_j \notin B_\delta(t_*), j = 1, \dots, m \}.$$

It follows that

$$M_\delta^{2m} \supset \phi_{el}^p(\Lambda^m \times T_\delta^m)$$

and therefore

$$Vol\left(M_\delta^{2m}\right) \geq Vol\left(\phi_{el}^p(\Lambda^m \times T_\delta^m)\right). \quad (10.1)$$

Using the identity $\int_{\Lambda^m} \prod_{i=1}^m \lambda_i d\lambda = \frac{1}{(2m)!}$, we compute the righthand side using the change of variables formula as

$$\begin{aligned} Vol\left(\phi_{el}^p(\Lambda^m \times T_\delta^m)\right) &= \int_{\Lambda^m \times T_\delta^m} |det(d\phi_{el}^p)| \\ &= \int_{\Lambda^m} \left(\prod_{j=1}^m \lambda_j\right) d\lambda \int_{T_\delta^m} \mathcal{J}_{el}^p \\ &= \frac{1}{(2m)!} \int_{T_\delta^m} \mathcal{J}_{el}^p \\ &= \frac{1}{(2m)!m!} \int_{I_\delta^m} \mathcal{J}_{el}^p \end{aligned}$$

where

$$I_\delta^m := \{(t_1, \dots, t_m) \in I^m : t_j \notin B_\delta, j = 1, \dots, m\}.$$

To bound this from below we bound the integral over $(I_\delta^m)^c$ from above. To that end, let

$$I_{\delta,j}^m := \{(t_1, \dots, t_m) \in I^m : t_j \in B_\delta\}, \quad j = 1, \dots, m,$$

so that

$$(I_\delta^m)^c = \cup_j I_{\delta,j}^m.$$

Therefore, using a union bound, we have

$$\begin{aligned}
\int_{(I_\delta^m)^c} \mathcal{J}_{el}^p &= \int_{(I_\delta^m)^c} \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t) dt \\
&= \int_{\cup_{j'=1}^m I_{\delta,j'}^m} \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t) dt \\
&\leq \sum_{j'=1}^m \int_{I_{\delta,j'}^m} \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t) dt \\
&= m \int_{I_{\delta,1}^m} \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t) dt \\
&= m \int_{I_{\delta,1}^m} \prod_{j=1}^m t_j^2 \prod_{1 \leq j < k \leq m} (t_k - t_j)^4 dt_1 \cdots dt_m \\
&\leq m \int_{I_{\delta,1}^m} \prod_{j=2}^m t_j^2 \prod_{2 \leq j < k \leq m} (t_k - t_j)^4 dt_1 \cdots dt_m \\
&= m \text{Vol}(B_\delta) \int_{I^{m-1}} \prod_{j=2}^m t_j^2 \cdot \Delta_{m-1}^4(t) dt \\
&= m \text{Vol}(B_\delta) S_{m-1}(3, 1, 2) \\
&\leq 2m\delta S_{m-1}(3, 1, 2)
\end{aligned}$$

and so obtain

$$\begin{aligned}
\frac{\text{Vol}\left(\phi_{el}^p(\Lambda^m \times T_\delta^m)\right)}{\text{Vol}\left(\phi_{el}^p(\Lambda^m \times T^m)\right)} &= \frac{\int_{I_\delta^m} \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t) dt}{\int_{I^m} \prod_{j=1}^m t_j^2 \cdot \Delta_m^4(t) dt} \\
&\geq 1 - 2m\delta \frac{S_{m-1}(3, 1, 2)}{S_m(3, 1, 2)}
\end{aligned}$$

Using Selberg's formulas (6.1) we compute

$$\begin{aligned}
\frac{S_{m-1}(3, 1, 2)}{S_m(3, 1, 2)} &= \frac{\prod_{j=0}^{m-2} \frac{\Gamma(1+2j)\Gamma(3+2j)^2}{2\Gamma(2(m+j))}}{\prod_{j=0}^{m-1} \frac{\Gamma(1+2j)\Gamma(3+2j)^2}{2\Gamma(2(m+j)+2)}} \\
&= \frac{2\Gamma(4m-2)}{\Gamma(2m-1)^2\Gamma(2m+1)} \frac{\prod_{j=0}^{m-1} \frac{\Gamma(1+2j)\Gamma(3+2j)^2}{2\Gamma(2(m+j))}}{\prod_{j=0}^{m-1} \frac{\Gamma(1+2j)\Gamma(3+2j)^2}{2\Gamma(2(m+j)+2)}} \\
&= \frac{2\Gamma(4m-2)}{\Gamma(2m-1)^2\Gamma(2m+1)} \prod_{j=0}^{m-1} \frac{\Gamma(2(m+j)+2)}{\Gamma(2(m+j))} \\
&= \frac{2\Gamma(4m-2)}{\Gamma(2m-1)^2\Gamma(2m+1)} \frac{\Gamma(4m)}{\Gamma(2m)}.
\end{aligned}$$

10.2 Even case of Proof of Theorem 7.1

For the even case, $n = 2m$, recall the identity $\mathcal{J}_{eu}^c(0, t) \equiv 0$, and the volume identity (6.8)

$$Vol(M^{2m}) = \frac{1}{(2m)!m!} \int_{I^m} \mathcal{J}_{el}^c(t_*, t) dt + \frac{1}{(2m)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt \quad (10.2)$$

Now observe that, for $i \geq 0$, the identity

$$\begin{aligned}
&\int_{M^{2m}} q_i - \frac{t_*^i}{2m+1} Vol(M^{2m}) \\
&= \frac{\delta_0(i) + 1}{(2m+1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt \\
&+ \frac{2}{(2m+1)!m!} \int_{I^m} \Sigma t^i \mathcal{J}_{el}^c(t_*, t) dt + \frac{2}{(2m+1)!(m-1)!} \int_{I^{m-1}} \Sigma t^i \mathcal{J}_{eu}^c(t_*, t) dt
\end{aligned}$$

from Proposition 6.1, evaluated at $t_* = 0$ becomes

$$\begin{aligned}
&\int_{M^{2m}} q_i - \frac{\delta_0(i)}{2m+1} Vol(M^{2m}) \\
&= \frac{\delta_0(i) + 1}{(2m+1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(0, t) dt \\
&+ \frac{2}{(2m+1)!m!} \int_{I^m} \Sigma t^i \mathcal{J}_{el}^c(0, t) dt + \frac{2}{(2m+1)!(m-1)!} \int_{I^{m-1}} \Sigma t^i \mathcal{J}_{eu}^c(0, t) dt
\end{aligned}$$

where we note that $\Sigma t^i = \sum_{j=1}^m t_j^i$ in the integral over I^m and $\Sigma t^i = \sum_{j=1}^{m-1} t_j^i$ in the integral over I^{m-1} .

Subtracting the two, using the identity $\mathcal{J}_{eu}^c(0, t) \equiv 0$, we obtain

$$\begin{aligned}
& -\frac{t_*^i}{2m+1} \text{Vol}(M^{2m}) + \frac{\delta_0(i)}{2m+1} \text{Vol}(M^{2m}) \\
= & \frac{\delta_0(i) + 1}{(2m+1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt \\
+ & \frac{2}{(2m+1)!m!} \int_{I^m} \Sigma t^i (\mathcal{J}_{el}^c(t_*, t) - \mathcal{J}_{el}^c(0, t)) dt + \frac{2}{(2m+1)!(m-1)!} \int_{I^{m-1}} \Sigma t^i \mathcal{J}_{eu}^c(t_*, t) dt
\end{aligned}$$

and applying the volume identity (10.2) we obtain

$$\begin{aligned}
& -\frac{t_*^i}{2m+1} \text{Vol}(M^{2m}) \\
= & -\frac{\delta_0(i)}{(2m+1)!m!} \int_{I^m} \mathcal{J}_{el}^c(t_*, t) dt + \frac{1}{(2m+1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt \\
+ & \frac{2}{(2m+1)!m!} \int_{I^m} \Sigma t^i (\mathcal{J}_{el}^c(t_*, t) - \mathcal{J}_{el}^c(0, t)) dt + \frac{2}{(2m+1)!(m-1)!} \int_{I^{m-1}} \Sigma t^i \mathcal{J}_{eu}^c(t_*, t) dt.
\end{aligned}$$

Then, if we let $\phi(s) := \sum_{i=0}^{2m} \phi_i s^i$ be a polynomial of degree $n = 2m$, summing over each identity in (10.3), we conclude that

$$\begin{aligned}
& -\frac{\phi(t_*)}{2m+1} \text{Vol}(M^{2m}) \\
= & \phi_0 \left(\frac{1}{(2m+1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt - \frac{1}{(2m+1)!m!} \int_{I^m} \mathcal{J}_{el}^c(t_*, t) dt \right) \\
+ & \sum_{j=1}^{2m} \phi_j \frac{1}{(2m+1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt \\
+ & \frac{2}{(2m+1)!m!} \int_{I^m} (\Sigma \phi)(t) (\mathcal{J}_{el}^c(t_*, t) - \mathcal{J}_{el}^c(0, t)) dt + \frac{2}{(2m+1)!(m-1)!} \int_{I^{m-1}} (\Sigma \phi)(t) \mathcal{J}_{eu}^c(t_*, t) dt \\
= & -\phi_0 \frac{1}{(2m+1)!m!} \int_{I^m} \mathcal{J}_{el}^c(t_*, t) dt \\
+ & \sum_{j=0}^{2m} \phi_j \frac{1}{(2m+1)!(m-1)!} \int_{I^{m-1}} \mathcal{J}_{eu}^c(t_*, t) dt \\
+ & \frac{2}{(2m+1)!m!} \int_{I^m} (\Sigma \phi)(t) (\mathcal{J}_{el}^c(t_*, t) - \mathcal{J}_{el}^c(0, t)) dt + \frac{2}{(2m+1)!(m-1)!} \int_{I^{m-1}} (\Sigma \phi)(t) \mathcal{J}_{eu}^c(t_*, t) dt
\end{aligned}$$

Since $\phi(0) = \phi_0$ and $\phi(1) = \sum_{i=0}^{2m} \phi_i$ the assertion follows by multiplication by $2m+1$.

10.3 Assorted Technical Results

Proposition 10.1. *We have*

$$m^2 \leq 8 \left(\frac{e}{2}\right)^{4m}, \quad m \geq 1.$$

Proof. We proceed by induction. The inequality is clearly true for $m = 1, 2$. Therefore, suppose that it is true for some $m \geq 2$. Then, since

$$\frac{m+1}{m} \leq \frac{3}{2} \leq \left(\frac{e}{2}\right)^2, \quad m \geq 2$$

follows from $6 \leq e^2$, we conclude that

$$\begin{aligned} (m+1)^2 &= \frac{(m+1)^2}{m^2} m^2 \\ &\leq \frac{(m+1)^2}{m^2} 8 \left(\frac{e}{2}\right)^{4m} \\ &\leq \left(\frac{e}{2}\right)^4 8 \left(\frac{e}{2}\right)^{4m} \\ &\leq 8 \left(\frac{e}{2}\right)^{4m+4} \end{aligned}$$

thus establishing the inequality for $m+1$ and finishing the proof. \square

Proposition 10.2. *We have*

$$B(a, a) \geq \frac{4}{a} 2^{-2a}, \quad a > 1$$

Proof. We have from the integral formula

$$\begin{aligned} B(a, b) &= \int_0^1 t^{a-1} (1-t)^{a-1} dt \\ &= 2 \int_0^{\frac{1}{2}} t^{a-1} (1-t)^{a-1} dt \\ &\geq 2 \left(\frac{1}{2}\right)^{a-1} \int_0^{\frac{1}{2}} t^{a-1} dt \\ &= 2 \left(\frac{1}{2}\right)^{a-1} \frac{1}{a} \left(\frac{1}{2}\right)^a \\ &= \frac{4}{a} 2^{-2a} \end{aligned}$$

\square

Acknowledgements

We would like to thank Gérard Letac for his helpful comments, in particular for his substantial simplification, included here, of our previous proof of Lemma 4.1. We would also like to thank one of the referees for many helpful comments which we also feel improved the manuscript.

The authors gratefully acknowledge this work supported by the Air Force Office of Scientific Research under Award Number FA9550-12-1-0389 (Scientific Computation of Optimal Statistical Estimators).

References

- [1] C. Abraham and B. Cadre. Asymptotic properties of posterior distributions derived from misspecified models. *C. R. Math. Acad. Sci. Paris*, 335(5):495–498, 2002.
- [2] C. Abraham and B. Cadre. Concentration of posterior distributions with misspecified models. *Ann. I.S.U.P.*, 52(3):3–14, 2008.
- [3] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1965.
- [4] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, Berlin, third edition, 2006.
- [5] G. B. Arfken and H. J. Weber. *Mathematical Methods for Physicists*. Elsevier Academic Press, Burlington, Mass., sixth edition, 2005.
- [6] R. R. Bahadur and L. J. Savage. The nonexistence of certain statistical procedures in nonparametric problems. *Ann. Math. Statist.*, 27(4):1115–1122, 1956.
- [7] S. N. Bernšteĭn. *Sobranie sochinenii. Tom IV: Teoriya veroyatnostei. Matematicheskaya statistika. 1911–1946*. Izdat. “Nauka”, Moscow, 1964.
- [8] D. Bertsimas and I. Popescu. Optimal inequalities in probability theory: a convex optimization approach. *SIAM J. Optim.*, 15(3):780–804 (electronic), 2005.
- [9] V. I. Bogachev. *Gaussian Measures*. Mathematical Surveys and Monographs, Vol. 62. American Mathematical Society, Providence, 1998.
- [10] C. Castaing and M. Valadier. *Convex Analysis and Measurable Multifunctions*. Lecture Notes in Mathematics, Vol. 580. Springer-Verlag, Berlin, 1977.
- [11] F-C. Chang, J. H. B. Kemperman, and W. J. Studden. A normal limit theorem for moment sequences. *Annals of Probability*, 21(3):1295–1309, 1993.
- [12] E. B. Christoffel. Über die Gaußsche Quadratur und eine Verallgemeinerung derselben. *J. Reine Angew. Math.*, 55:61–82, 1858.
- [13] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*. MIT Press, Cambridge, Mass., 1990.
- [14] P. Diaconis and D. Freedman. On the consistency of Bayes estimates. *Ann. Statist.*, 14(1):1–67, 1986. With a discussion and a rejoinder by the authors.
- [15] P. W. Diaconis and D. Freedman. Consistency of Bayes estimates for nonparametric regression: normal theory. *Bernoulli*, 4(4):411–444, 1998.
- [16] D. L. Donoho. One-sided inference about functionals of a density. *Ann. Statist.*, 16(4):1390–1420, 1988.

- [17] D. Draper. Assessment and propagation of model uncertainty. *J. Roy. Statist. Soc. Ser. B*, 57(1):45–97, 1995. With discussion and a reply by the author.
- [18] D. Draper. Bayesian model specification: heuristics and examples. In P. Damien, P. Dellaportas, N. G. Polson, and D. A. Stephens, editors, *Bayesian Theory and Applications*. Oxford University Press, 2013.
- [19] P. J. Forrester and S. O. Warnaar. The importance of the Selberg integral. *Bull Amer. Math. Soc.*, 45(4):489–534, 2008.
- [20] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, San Diego, seventh edition, 2007.
- [21] P. Grünwald and J. Langford. Suboptimal behavior of Bayes and MDL in classification under misspecification. In *Learning theory*, volume 3120 of *Lecture Notes in Comput. Sci.*, pages 331–347. Springer, Berlin, 2004.
- [22] Paul Gustafson. On measuring sensitivity to parametric model misspecification. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 63(1):81–94, 2001.
- [23] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.
- [24] S. Karlin and L. S. Shapley. *Geometry of Moment Spaces*. Memoirs of the American Mathematical Society, Number 12. American Math. Soc., 1953.
- [25] S. Karlin and W. J. Studden. *Tchebycheff Systems: With Applications in Analysis and Statistics*. Pure and Applied Mathematics, Vol. XV. Interscience Publishers John Wiley & Sons, New York-London-Sydney, 1966.
- [26] B. J. K. Kleijn and A. W. van der Vaart. The Bernstein-Von-Mises theorem under misspecification. *Electron. J. Stat.*, 6:354–381, 2012.
- [27] M. G. Kreĭn. The ideas of P. L. Čebyšev and A. A. Markov in the theory of limiting values of integrals and their further development. In E. B. Dynkin, editor, *Eleven papers on Analysis, Probability, and Topology*, American Mathematical Society Translations, Series 2, Volume 12, pages 1–122. American Mathematical Society, New York, 1959.
- [28] L. LeCam. On some asymptotic properties of maximum likelihood estimates and related Bayes’ estimates. *Univ. California Publ. Statist.*, 1:277–329, 1953.
- [29] H. Lian. On rates of convergence for posterior distributions under misspecification. *Comm. Statist. Theory Methods*, 38(11-12):1893–1900, 2009.
- [30] H. Owhadi, C. Scovel, and T. J. Sullivan. Bayesian Brittleness. *arXiv:1304.6772*, 2013.

- [31] H. Owhadi, C. Scovel, T. J. Sullivan, M. McKerns, and M. Ortiz. Optimal Uncertainty Quantification. *SIAM Review*, to appear, 2013. arXiv:1009.0679 (published online on 2010).
- [32] Grünwald P. Bayesian inconsistency under misspecification, 2006.
- [33] P. Ressel. Some continuity and measurability results on spaces of measures. *Mathematica Scandinavica*, 40:69–78, 1977.
- [34] R. T. Rockafellar. *Convex Analysis*. Princeton Univ. Press, Princeton, 1970.
- [35] M.-F. Sainte-Beuve. On the extension of von Neumann-Aumann’s theorem. *J. Functional Analysis*, 17:112–129, 1974.
- [36] S. Saitoh. *Integral transforms, reproducing kernels and their applications*. Addison Wesley Longman, Essex, 1997.
- [37] F. J. Samaniego. *A comparison of the Bayesian and frequentist approaches to estimation*. Springer Series in Statistics. Springer, New York, 2010. <http://dx.doi.org/10.1007/978-1-4419-5941-6>.
- [38] B. Simon. The Christoffel-Darboux kernel. In *Perspectives in PDE, Harmonic Analysis and Applications, a volume in honor of VG Maz’ya’s 70th birthday*, pages 295–335. American Math. Soc., 2008.
- [39] M. Skibinsky. The range of the $(n + 1)$ th moment for distributions on $[0, 1]$. *Journal of Applied Probability*, 4(3):543–552, 1967.
- [40] J. E. Smith. Generalized Chebychev inequalities: Theory and applications in decision analysis. *Operations Research*, 43(5):807–825, 1995.
- [41] M. Spivak. *Calculus on Manifolds*. Addison-Wesley, Reading, Massachusetts, 1965.
- [42] R. von Mises. *Mathematical Theory of Probability and Statistics*. Edited and Complemented by Hilda Geiringer. Academic Press, New York, 1964.
- [43] H. White. Maximum likelihood estimation of misspecified models. *Econometrica*, 50(1):1–25, 1982.